Quantum Information Processing

From Theory to Experiment

Edited by

Dimitris G. Angelakis

University of Cambridge, United Kingdom

Matthias Christandl

University of Cambridge, United Kingdom

Artur Ekert

University of Cambridge, United Kingdom and
National University of Singapore

Alastair Kay

University of Cambridge, United Kingdom

and

Sergei Kulik

Moscow M.V. Lomonosov State University, Russia

IOS Press

Amsterdam • Berlin • Oxford • Tokyo • Washington, DC

Published in cooperation with NATO Public Diplomacy Division

Proceedings of the NATO Advanced Study Institute on
Quantum Computation and Quantum Information
Chania, Crete, Greece
2–13 May 2005

© 2006 IOS Press.

All rights reserved. No part of this book may be reproduced, stored in a retrieval system,
or transmitted, in any form or by any means, without prior written permission from the publisher.

ISBN 1-58603-611-4
Library of Congress Control Number: 2006927263

Publisher
IOS Press
Nieuwe Hemweg 6B
1013 BG Amsterdam
Netherlands
fax: +31 20 687 0019
e-mail: order@iospress.nl

Distributor in the UK and Ireland
Gazelle Books Services Ltd.
White Cross Mills
Hightown
Lancaster LA1 4XS
United Kingdom
fax: +44 1524 63232
e-mail: sales@gazellebooks.co.uk

Distributor in the USA and Canada
IOS Press, Inc.
4502 Rachael Manor Drive
Fairfax, VA 22032
USA
fax: +1 703 323 3668
e-mail: iosbooks@iospress.com

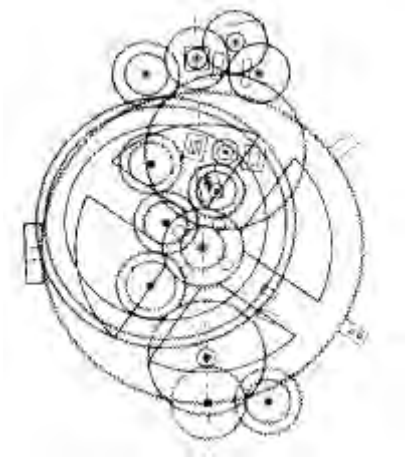
LEGAL NOTICE
The publisher is not responsible for the use which might be made of the following information.

PRINTED IN THE NETHERLANDS

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Introduction

By Artur EKERT



Chania, a picturesque little town on the coast of the western Crete, the birthplace of Dimitris Angelakis, the main organizer of this Advanced Study Institute in Quantum Information Science, is only a few miles away from the tiny island of Antikythera. In 1900 a party of sponge-divers were driven by a storm to anchor near the island and there, at a depth of some 40 meters, they found the wreck of an ancient cargo ship. Among the pieces of pottery and marble statutes sprawled on the seabed was a coral-encrusted lump of corroded bronze gear wheels. The Antikythera mechanism, as it is now known, was probably the world's first "analog computer" — a sophisticated device for calculating the motions of stars and planets. This remarkable assembly of more than 30 gears with a differential mechanism, made on Rhodes or Cos in the first century B.C., revised the view of what the ancient Greeks were capable of creating at that time. A comparable level of engineering didn't become widespread until the industrial revolution nearly two millennia later. Thus it is hardly surprising that Richard Feynman, who saw the Antikythera mechanism on display in Athens, called it "nearly impossible". In one of his letters, reprinted in "What Do You Care What Other People Think?: Further Adventures of a Curious Character", he wrote

"*... Yesterday morning I went to the archaeological museum. ... Also, it was slightly boring because we have seen so much of that stuff before. Except for one thing: among all those art objects there was one thing so entirely different and strange that it is nearly impossible. It was recovered from the sea in 1900 and is some kind of machine with gear trains, very much like the inside of a modern wind-up alarm clock. The teeth are very regular and many wheels are fitted closely together...*"

It seems likely that the Antikythera tradition of complex mechanical technology was transmitted via the Arab world to medieval Europe where it formed the basis of clockmaking techniques. As such, the Antikythera mechanism is a venerable precursor of mechanical computing devices based on the meshing of metal gears. Indeed, for many years the basic raw material of the computer industry was brass. The 17th century calculators of Wilhelm Schickard, Blaise Pascal and Gottfried Wilhelm Leibniz testify to the importance of gears in the history of computing. When, in 1837, Charles Babbage was tinkering with a design of the first programmable computer, known as the Analytical Engine, he was thinking in terms of rods, gears and wheels.

From gears to relays to valves to transistors to integrated circuits and so on – in the 20th century brass gave way to silicon. Today's advanced lithographic techniques can etch logic gates and wires less than a micron across onto the surfaces of silicon chips. Soon they will yield even smaller components, until we reach the point where logic gates are so small that they consist of only a few atoms each. If computers are to continue to become faster (and therefore smaller), new, quantum technology must replace or supplement what we have now, but it turns out that such technology can offer much more than smaller and faster microprocessors. It can support entirely new modes of computation, with new quantum algorithms that do not have classical analogues.

The very same person who was so fascinated by the ancient Antikythera laid down the foundations of quantum computation. In 1981 Feynman observed that simulations of some quantum experiments on any classical computer appear to involve an exponential slowdown in time as compared to the natural run of the experiment. Instead of viewing this fact as an obstacle, Feynman regarded it as an opportunity. If it requires so much computation to work out what will happen in a complicated quantum experiment then, he argued, the very act of setting up an experiment and measuring the outcome is tantamount to performing a complex computation. After all, any real computation is a physical process, be it classical or quantum. Thus any computation can be viewed in terms of physical experiments which produce outputs that depend on initial preparations called inputs. Since then, the hunt has been on for interesting things for quantum computers to do, and at the same time, for the scientific and technological advances that could allow us to build quantum computers.

The NATO Advanced Study Institute in Chania brought together a number of researchers and students in both experimental and theoretical quantum information science. During lectures and talks, and in numerous discussions over Raki, the participants shared their views on just about everything; from quantum algorithms and intricacies of computational complexity to the finer parts of Cretan cuisine, and from new technologies for realizing quantum computers to the spirit of traditional Greek dances. The knowledge that nature can be coherently controlled and manipulated at the quantum level was perceived as both a powerful stimulus and one of the greatest challenges facing experimental physics. Fortunately the exploration of quantum technology has many staging posts along the way, each of which will yield scientifically and technologically useful results and some of them are described in this volume.

We hope this collection of papers provides a good overview of the current state-of-the-art of quantum information science. We do not know how a quantum Antikythera will look like but all we know is that the best way to predict the future is to create it. From the perspective of the future, it may well be that the real computer age has not yet even begun.

We also wish to thank our sponsors NATO, the Cambridge-MIT Institute, the Union of Agricultural Cooperatives of Kidonia and Kissamos, the Cooperative Bank of Chania, ANEK Lines, Olympic Airways, ABEA Olive Oil Products and Stigmes Magazine. Finally we acknowledge the helpful collaboration from IOS Press in the publication of this volume and also thank Kaija Hampson for being an excellent secretary during the meeting.

Contents

Introduction v
Artur Ekert

Chapter 1. Quantum Communication and Entanglement

Quantum Entanglement: Detection Methods and Usefulness as a Physical Resource 3
Dagmar Bruß

On Quantum Cryptography with Bipartite Bound Entangled States 19
Paweł Horodecki and Remigiusz Augusiak

Unitary Local Permutations on Bell Diagonal States of Qudits and Quantum Distillation Protocols 30
Hector Bombin and Miguel A. Martin-Delgado

Quantum Communication Channels in Infinite Dimensions 41
Alexander S. Holevo

Introduction to Relativistic Quantum Information 61
Daniel Terno

Generalized Bell Inequalities and the Entanglement of Pure States 83
Kwek Leong-Chuan, Chunfeng Wu, Jingling Chen, Dagomir Kaszlikowski and C.H. Oh

Thermal Entanglement in Infinite Dimensional Systems 89
Aires Ferreira, Ariel Guerreiro and Vlatko Vedral

Improved Algorithm for Quantum Separability and Entanglement Detection 93
Lawrence M. Ioannou, Benjamin C. Travaglione, Donny Cheung and Artur K. Ekert

Generalised Entanglement Swapping 99
Anthony J. Short, Sandu Popescu and Nicolas Gisin

Quantum Information Processing with Low-Dimensional Systems 103
Alexander Yu. Vlasov

Local Information and Nonorthogonal States 109
Jonathan Walgate

Optimal Alphabets for Noise-Resistant Quantum Cryptography 113
Denis Sych, Boris Grishanin and Victor Zadkov

Chapter 2. Quantum Algorithms and Error Correction

Quantum Algorithms and Complexity 121
Michele Mosca

An Introduction to Measurement Based Quantum Computation 137
Richard Jozsa

Quantum Error Correction and Fault-Tolerance 159
Daniel Gottesman

Simulating Fourier Transforms for Open Quantum Systems in Higher Encoding Bases 170
Ioannis N. Doxaras

Entanglement, Area Law and Group Theory 175
Radu Ionicioiu, Alioscia Hamma and Paolo Zanardi

A Quantum Algorithm for Closest Pattern Matching 180
Paulo Mateus and Yasser Omar

Classical and Quantum Fingerprinting in the One-Way Communication Model 184
Anthony J. Scott, Jonathan Walgate and Barry C. Sanders

Chapter 3. Quantum Information Theory in Spin Systems

Introduction to Localizable Entanglement 191
Markus Popp, Frank Verstraete, Miguel A. Martin-Delgado and Ignacio Cirac

State Transfer in Permanently Coupled Quantum Chains 218
Daniel Burgarth, Vittorio Giovannetti and Sougato Bose

Geometric Effects in Spin Chains 238
Marie Ericsson and Alastair Kay

Quantum Walks and Decoherence on a 1+1 Lattice 242
Isabelle Herbauts

Certain Aspects of Quantum Random Walk Asymptotics 246
Ioannis Smyrnakis

Chapter 4. Implementations of Quantum Information Processing

From Entanglement to Quantum Key Distribution 255
Hannes Hübel and Anton Zeilinger

Preparation and Measurement of Qutrits Based on Single-Mode Biphotons 281
Sergei Kulik

Cavity Quantum Electrodynamics: Quantum Entanglement and Information 294
Jean-Michel Raimond

Possibility of Quantum Computation by Utilizing Carbon Nanotubes – Cooper Pair Splitting by Tomonaga-Luttinger Liquid – 312
Junji Haruyama, K. Murakami, J. Mizubayashi and N. Kobayashi

Implementing Quantum Processors in the Solid State 321
Crispin H.W. Barnes

Quantum Field Trajectories Under Photon Number QND Measurement 326
Alexander A. Bukach and Sergei Ya. Kilin

Quantum Computation Beyond the "Standard Circuit Model" 330
Konstantinos Ch. Chatzisavvas, Costas Daskaloyannis and Christos P. Panos

Entangled Light from Optical Time Boundaries 337
Ariel Guerreiro, Aires Ferreira, José T. Mendonça and Vlatko Vedral

Strong Light-Matter Coupling: Parametric Interactions in a Cavity and Free-Space 341
Igor B. Mekhov, Valentin S. Egorov, Victor N. Lebedev, Peter V. Moroshkin, Igor A. Chekhonin and Sergei N. Bagayev

Macroscopic Quantum Information Channel Via the Polarization-Sensitive Interaction Between the Light and Spin Subsystems 346
Oksana S. Mishina, Dmitriy V. Kupriyanov and Eugene S. Polzik

From Network Complexity to Time Complexity Via Optimal Control 353
Thomas Schulte-Herbrüggen, Andreas Spörl, Navin Khaneja and Steffen Glaser

Author Index 359

This page intentionally left blank

Chapter 1

Quantum Communication and Entanglement

This page intentionally left blank

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Quantum Entanglement: Detection Methods and Usefulness as a Physical Resource

Dagmar Bruß
Institute for Theoretical Physics, Universität Düsseldorf, Germany

Abstract. Entanglement is one of the most fascinating features of quantum information. In this lecture the focus is on two aspects of the broad field of entanglement theory: first, on methods for the detection or verification of entanglement. Here, the method of Bell inequalities is compared with the tool of witness operators and their local measurement. An overview of other methods such as entropic and uncertainty relations, as well as physical approximations of unphysical maps, will be given. Second, the zoo of multipartite entangled states will be visited, mainly with respect to their usefulness as a resource for a given task. As an example, distributed dense coding is presented.

As a connection between the two main topics, the detection of multipartite entangled states is discussed.

Keywords. Quantum entanglement, Bell inequalities, witness operators, uncertainty relations, physical maps, multipartite entanglement

1. Detecting Entanglement

1.1. Bipartite Entanglement

As a reminder, let us recall the definition of bipartite entanglement for pure states. We consider a pure state which is an element of a composite Hilbert space of two subsystems, namely $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. We call $|\psi\rangle$ separable, iff

$$|\psi\rangle = |a\rangle \otimes |b\rangle, \tag{1}$$

and entangled otherwise.

For mixed states, the definition of a separable state goes back to [1], and reads: ϱ is separable iff

$$\varrho = \sum_i p_i |a_i\rangle\langle a_i| \otimes |b_i\rangle\langle b_i|, \tag{2}$$

with $0 \leq p_i \leq 1$, $\sum_i p_i = 1$, i.e. iff the density operator can be written as a convex combination of projectors onto product states. Otherwise ϱ is called

entangled. Note that eq. (2) is *not* a spectral decomposition, and that in general $\langle a_i|\, a_j\rangle \neq \delta_{ij}, \langle b_i|\, b_j\rangle \neq \delta_{ij}$ may hold.

There exist several operational separability criteria in the literature, for overviews see [2]. For pure states a simple tool to find out whether a state is entangled is the Schmidt decomposition. It states that any bipartite state can be written as a bi-orthogonal superposition,

$$|\,\psi^r\rangle = \sum_i^r a_i|\,e_i\rangle|\,f_i\rangle, \tag{3}$$

where the Schmidt coefficients a_i can be chosen positive, $a_i > 0$, and are normalised, $\sum_i^r a_i^2 = 1$. The orthogonality condition reads $\langle e_i|\, e_j\rangle = \delta_{ij} = \langle f_i|\, f_j\rangle$. The Schmidt rank r has to be smaller or equal to the minimum of the two dimensions of the two Hilbert spaces. The state $|\,\psi\rangle$ is separable iff $r = 1$. The Schmidt rank can be easily determined by taking the partial trace over one of the subsystems.

The most famous separability criterion for mixed states is the Peres-Horodecki criterion [3,4], or criterion of positive partial transpose (PPT). The partial transpose is defined as the transpose of one of the two subsystems, e.g. the partial transpose for Alice's system reads

$$(\varrho^{T_A})_{m\mu,n\nu} = \varrho_{n\mu,m\nu}. \tag{4}$$

Asher Peres realised that separability implies positivity of the partial transpose. This can be seen directly from the definition of a separable state in eq. (2). The reverse only holds in low dimensions, namely in dimensions 2×2 and 2×3, as pointed out by the Horodecki family, who used properties of positive maps to prove this statement.

1.2. Detecting Entanglement Via Bell Inequalities

One possibility to detect experimentally whether a given state is entangled relies on a Bell inequality test. An example of a Bell inequality is the so-called CHSH-inequality [5]. We assume a, b, c, d to be classical variables that can take the values ± 1. Thus, $(a + c)b + (-a + c)d = \pm 2$. If we consider several assignments for the variables, and average over them, we arrive at

$$S = |\langle ab\rangle + \langle bc\rangle + \langle cd\rangle - \langle da\rangle| \leq 2, \tag{5}$$

which is the CHSH inequality.

In quantum mechanics, we can consider observables that lead to assignments to the variables. Explicitly, we can make a spin measurement in a specified direction. We thus arrive at $\langle ab\rangle = \mathrm{Tr}[(\vec{\alpha}\cdot\vec{\sigma}_A)\otimes(\vec{\beta}\cdot\vec{\sigma}_B)\varrho]$, where $\vec{\sigma}_{A(B)}$ denotes the vector of Pauli matrices for Alice (Bob), and $\vec{\alpha}(\vec{\beta})$ denotes a measurement direction for Alice (Bob). By choosing appropriate measurement directions, as indicated in figure 1, we find for the singlet $|\,\psi^-\rangle = \frac{1}{\sqrt{2}}(|\,01\rangle - |\,10\rangle)$ that $S = 2\sqrt{2} \geq 2$, a

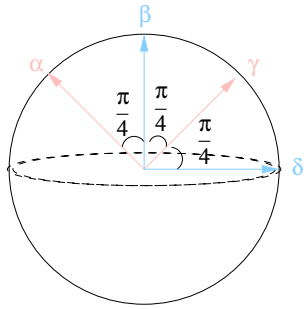


Figure 1. Optimal measurement directions on the Bloch sphere for testing the CHSH inequality.

clear indication of a violation of a CHSH inequality. This result is not compatible with a local hidden variable model, and in particular it shows that the structure of the singlet cannot be explained by classical correlations. However, as pointed out by Werner [1], there exist mixed entangled states of two qubits that do not violate any Bell inequality. Therefore, the tool of Bell inequalities is not the best method to detect entanglement. This is not astonishing, as it was not designed for entanglement detection, but for excluding local hidden variable models, i.e. starting from purely classical arguments.

1.3. Detecting Entanglement Via Witness Operators

When designing a tool for the detection of quantum entanglement, it is advantageous to start from the structure of quantum states. An important property of quantum states is convexity: mixing two density operators leads to another density operator. Separable states also form a convex set: from the definition in eq. (2) it is clear that mixing two separable density operators (i.e. two sums of projectors onto product states) leads to another separable density operator. The concept of witness operators is based on this convex structure of quantum states, which is illustrated in figure 2.

Entanglement witnesses were introduced in [4,6] and provide a necessary and sufficient criterion for entanglement: ϱ is entangled iff there exists a Hermitean operator $\mathcal{W}$ with

$$\mathrm{Tr}(\mathcal{W}\varrho) < 0 \tag{6}$$

and $\mathrm{Tr}(\mathcal{W}\varrho_{sep}) \geq 0$ for all separable states ϱ_{sep}.

The optimisation of witnesses has been studied (a witness $\mathcal{W}$ is called optimal if there is no witness that detects all the states that are detected by $\mathcal{W}$, and some more in addition), and generic forms of witnesses have been constructed [7]. An example for a witness for a state ϱ with non-positive partial transpose is given by

$$\mathcal{W} = (|\,\phi\rangle\langle\phi\,|)^{T_A} \ , \tag{7}$$

with $\varrho^{T_A}|\,\phi\rangle = \lambda_{min}|\,\phi\rangle$, where λ_{min} denotes the minimal (negative) eigenvalue of the partially transposed state. It is straightforward to see that $\mathcal{W}$ is an entangle-

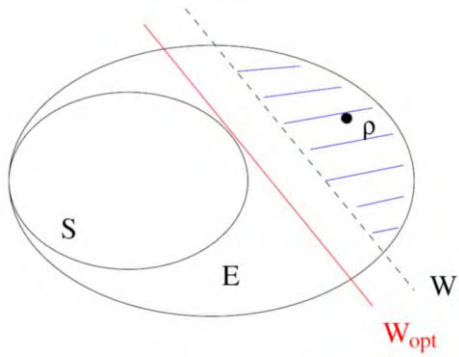


Figure 2. Set of all states: convex subset of separable states (S), entangled states (E), hyperplanes defined by witness W, and by optimised witness W_{opt}. The hatched area (in particular the state ρ) is detected by W.

ment witness that detects ϱ, as $Tr(\mathcal{W}\varrho) = Tr((|\,\phi\rangle\langle\phi\,|)^{T_A}\varrho) = Tr(|\,\phi\rangle\langle\phi\,|\varrho^{T_A}) = \langle\phi\,|\varrho^{T_A}|\,\phi\rangle = \lambda_{min} < 0$ and $Tr(\mathcal{W}\varrho_{sep}) = Tr((|\,\phi\rangle\langle\phi\,|)^{T_A}\varrho_{sep}) = Tr(|\,\phi\rangle\langle\phi\,|\varrho_{sep}^{T_A}) = \langle\phi\,|\varrho_{sep}^{T_A}|\,\phi\rangle \geq 0$. Another example for an entanglement witness is a typical construction that has been widely used in entanglement detection: given an entangled pure state $|\,\psi\rangle$,

$$\mathcal{W}_{|\,\psi\rangle} = x \cdot \mathbb{1} - |\,\psi\rangle\langle\psi\,| \tag{8}$$

with $x = \max_{|\,\varphi_{sep}\rangle} |\,\langle\psi|\,\varphi_{sep}\rangle|^2$ is an optimal entanglement witness that detects $|\,\psi\rangle\langle\psi\,|$. Again, this statement can be easily verified: $Tr(\mathcal{W}|\,\psi\rangle\langle\psi\,|) = x - |\,\langle\psi|\,\psi\rangle|^2 = x - 1 < 0$ and $Tr(\mathcal{W}\varrho_{sep}) = x - \langle\psi\,|\varrho_{sep}|\,\psi\rangle \geq 0$. We observe that $\mathcal{W}$ also detects $|\,\psi\rangle\langle\psi\,|$ plus some noise, as the noise only shifts the expectation value towards zero, as long as it is small enough. The problem of how to find x has been solved in [8]. One writes $|\,\psi\rangle$ as a certain bipartite superposition, i.e. $|\,\psi\rangle = \sum_{ij} C_{ij}|\,i\rangle|\,j\rangle$. A product state in the same bipartite split reads $|\,\varphi_{sep}\rangle = |\,A\rangle|\,B\rangle$, with $|\,A\rangle = \sum_i A_i|\,i\rangle$, $|\,B\rangle = \sum_j B_j|\,j\rangle$. Then $\tilde{x} = \sup_{A,B} |\sum_{ij} A_i^* B_j^* C_{ij}|^2 = \max \lambda(C)$, for the given bipartite split. In words, $\tilde{x}$ is the maximal (squared) Schmidt coefficient in a given bipartite split, and x is then the maximum of $\tilde{x}$ over *all* bipartite splits.

How can one use the mathematical concept of witness operators for the experimental detection of entanglement? In principle, one just has to determine the expectation value $Tr(\mathcal{W}\varrho)$. If this is negative, this is a clear signature of entanglement. However, how to measure $\mathcal{W}$? In the spectral decomposition one will find at least one entangled state. An important step for the implementation of the witness is to realise that *any* observable can be decomposed locally, and then measured locally with present-day technology. A local decomposition of a witness is not unique. Therefore, in a further step one can find the *optimal* local decomposition of a witness, where optimality here refers to the minimal number of measurement settings. This concept has been developed in [9].

Let us note that a local decomposition of the witness

$$\mathcal{W} = \sum_i c_i |\, a_i\rangle\langle a_i \,| \otimes |\, b_i\rangle\langle b_i \,| \;, \quad c_i \in \mathbb{R} \;, \quad \sum_i c_i = 1 \;, \tag{9}$$

has to be a so-called pseudo mixture, where at least one of the coefficients c_i is negative. This is due to the fact that a witness operator is necessarily non-positive.

As a simple example, let us consider a Bell state of two qubits plus white noise, namely $\rho = p|\,\psi^+\rangle\langle\psi^+\,| + (1-p)\mathbb{1}/4$, where $|\,\psi^+\rangle = \frac{1}{\sqrt{2}}(|\,01\rangle + |\,10\rangle)$, and p is a probability.

A suitable witness that is constructed according to the first method explained above is of the form $\mathcal{W} = (|\,\phi\rangle\langle\phi\,|)^{T_A}$ with $\rho^{T_A}|\,\phi\rangle = \lambda_{min}|\,\phi\rangle$, where here $|\,\phi\rangle = \alpha|00\rangle + \beta|11\rangle$ and α, β are parameters with $|\,\alpha\,|^2 + |\,\beta\,|^2 = 1$, which can actually be taken to be real.

It can be shown [9] that the optimal number of measurement settings for this case is three. The optimal local decomposition reads

$$\begin{aligned}
\mathcal{W} &= \alpha^2|\,z^+z^+\rangle\langle z^+z^+\,| + \beta^2|\,z^-z^-\rangle\langle z^-z^-\,| \\
&\quad +\alpha\beta(|\,x^+x^+\rangle\langle x^+x^+\,| + |\,x^-x^-\rangle\langle x^-x^-\,| \\
&\quad -|\,y^+y^-\rangle\langle y^+y^-\,| - |\,y^-y^+\rangle\langle y^-y^+\,|) \qquad (10) \\
&= \frac{1}{4}(\mathbb{1}\otimes\mathbb{1} + \sigma_z\otimes\sigma_z + (\alpha^2-\beta^2)(\sigma_z\otimes\mathbb{1} + \mathbb{1}\otimes\sigma_z) \\
&\quad +2\alpha\beta(\sigma_x\otimes\sigma_x + \sigma_y\otimes\sigma_y)) \qquad (11)
\end{aligned}$$

where $|z^\pm\rangle = |0\rangle, |1\rangle$; $|x^\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ and $|y^\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$.

Witnesses of this type were first measured in the experiment by F. DeMartini and collaborators [10]. Note that for the two-qubit case considered here one could also just perform state tomography, i.e. measure all elements of the density matrix. By applying the partial transpose to the reconstructed density matrix one can determine whether the produced state was entangled or not. In higher dimensions or for more than two subsystems this is in general not possible, because there exist bound entangled states which have a positive partial transpose. Therefore, the tool of witness operators becomes more powerful for the detection of entanglement than any previously known concept, when higher-dimensional and/or multipartite systems are considered, see subsection 2.4.

Note that recently the two concepts of Bell inequalities (which can always be reformulated as a non-optimal witness) and entanglement witnesses (which detect more states as entangled than Bell inequalities, and are thus more powerful) has been compared in [11].

1.4. Detecting Entanglement Via Uncertainty Relations

Recently, the method of using uncertainty relations for the detection of entanglement has been developed. There are two main ideas in this direction, namely variance based uncertainty relations and entropic uncertainty relations.

1.4.1. Variance Based Uncertainty Relations

Variance based uncertainty realtions for *local* observables were introduced by Hofmann and Takeuchi [12] and generalised to *non-local* observables by Gühne [13].

The variance of an observable M is defined as

$$\delta^2(M)_\varrho = \langle (M - \langle M \rangle_\varrho)^2 \rangle = \langle M^2 \rangle_\varrho - \langle M \rangle_\varrho^2 \ , \tag{12}$$

where $\langle M \rangle_\varrho = \mathrm{Tr}(\varrho M)$. A well-known uncertainty relation for non-commuting observables P and X reads $\delta(P)\delta(X) \geq \hbar/2$. A general formulation of an uncertainty relation for non-commuting observables M_i is given as

$$\sum_i \delta^2(M_i) \geq C > 0 \ . \tag{13}$$

In general, however, it is difficult to find the constant C.

The method for detection of entanglement consists in choosing $\{M_i\}$, finding a lower bound C for all product states, and then using concavity of the variance. Let us first show that the variance is concave with respect to a decomposition $\varrho = \sum_k p_k \varrho_k$:

$$\begin{aligned}
\delta^2(M_i)_\varrho &= \sum_k p_k \langle (M_i - \langle M_i \rangle_\varrho)^2 \rangle_{\varrho_k} \\
&= \sum_k p_k \left[\langle M_i^2 \rangle_{\varrho_k} - \langle M_i \rangle_{\varrho_k}^2 + \langle M_i \rangle_{\varrho_k}^2 - 2 \langle M_i \rangle_{\varrho_k} \langle M_i \rangle_\varrho + \langle M_i \rangle_\varrho^2 \right] \\
&= \sum_k p_k \left[\delta^2(M_i)_{\varrho_k} + (\langle M_i \rangle_{\varrho_k} - \langle M_i \rangle_\varrho)^2 \right] \\
&\geq \sum_k p_k \, \delta^2(M_i)_{\varrho_k} \ ,
\end{aligned} \tag{14}$$

and therefore

$$\sum_i \delta^2(M_i)_\varrho \geq \sum_k p_k \sum_i \delta^2(M_i)_{\varrho_k} \ . \tag{15}$$

Concavity of the variance means that one cannot decrease the uncertainty by mixing states. Thus, once having established a lower bound C of the variance for product states, a state with $\sum_i \delta^2(M_i) < C$ is entangled.

The local observables introduced in [12] read

$$M_i = A_i \otimes \mathbb{1} + \mathbb{1} \otimes B_i \ . \tag{16}$$

The local variances have lower bounds $\sum_i \delta^2(A_i)_{\varrho_A} \geq C_A$ and $\sum_i \delta^2(B_i)_{\varrho_B} \geq C_B$. It was shown that

$$\sum_i \delta^2(M_i)_\varrho \geq C_A + C_B \ , \tag{17}$$

when ϱ is separable. One problem of this approach is the generalisation to the multipartite scenario: no local uncertainty relation for the detection of multipartite entanglement is known yet.

In the approach of [13] non-local observables were introduced. It was shown that for any entangled pure bipartite state $|\psi\rangle$, there exists a set of observables $\{M_i\}$ such that $|\psi\rangle$ violates an uncertainty relation. As an example for the case that one subsystem is two-dimensional, i.e. $|\psi\rangle = a|00\rangle + b|11\rangle$ in the Schmidt decomposition, one can choose $M_i = |\psi_i\rangle\langle\psi_i|, i = 1...4$, with $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, |\psi_4\rangle\} = \{|\psi\rangle, |\psi\rangle^\perp, |\psi_3\rangle = a|01\rangle + b|10\rangle, |\psi_3\rangle^\perp\}$, where $|\psi\rangle^\perp$ denotes the state orthogonal to $|\psi\rangle$. It can be shown that for separable states $\sum_i \delta^2(M_i) \geq 2a^2b^2$. However, for the state $|\psi\rangle$ one finds $\sum_i \delta^2(M_i) = 0$, which proves that $|\psi\rangle$ is entangled, unless $a = 0$ or $b = 0$.

It is important to note that non-local measurements can be decomposed locally [9], and thus the non-locality of the observables is not a practical problem. For the detection of bound entangled states, e.g. the UPB states, one can find non-local measurement bases such that $\sum_i \delta^2(M_i) = 0$. Furthermore, multipartite entanglement can be detected via a generalisation of the ideas described above.

1.4.2. Entropic Uncertainty Relations

Entropic uncertainty relations were discussed in [14,15,16]. Again, one measures observables M_i. This leads to a probability distribution, with an according entropy $S(M_i)$. An entropic uncertainty relation for non-commuting observables M_i reads

$$\sum_i S(M_i) \geq C > 0 \; . \tag{18}$$

Note that the difference of this entanglement criterion to other entropic entanglement criteria such as the reduction criterion lies in the fact that there the eigenvalues of ϱ (i.e. outcomes of a measurement in the eigenbasis) are employed, whereas here the outcomes of more general measurements are used.

The idea for the detection of entanglement via entropic uncertainty relations is similar to the variance based uncertainty relations discusssed in 1.4.1: one chooses observables $\{M_i\}$, finds a lower bound C for all product states, and uses concavity of the entropy, which leads to the fact that a state with $\sum_i S(M_i) < C$ is detected as entangled.

An important open question is which entropy is the most suitable for the detection of entanglement. (In the literature the Tsallis entropy has been widely used.) Let us remark that the concept of entropic uncertainty relations is extendible to the detection of multipartite entanglement.

1.5. Physical Approximations of Positive Maps

A completely different idea for the detection of entanglement are physical approximations of positive maps [17]. Let us remember the separability criterion based on positive, but not completely positive (CP) maps: a bipartite state ϱ is separable iff $[\mathbb{1} \otimes \Lambda](\varrho) \geq 0$ for all positive maps Λ. For the detection of entanglement only those maps are interesting which are not completely positive, i.e. an

extension of which in some higher-dimensional space is not positive. Maps which are not CP are considered to be unphysical.

The idea in [17] is to add a suitable proportion of the identity map to $[\mathbb{1}\otimes\Lambda]$, such that the new map is positive, i.e. a physical map. This physical map for a $d\times d$ system is given as

$$\widetilde{[\mathbb{1}\otimes\Lambda]}(\varrho) = p\frac{[\mathbb{1}\otimes\mathbb{1}]}{d^2}(\varrho) + (1-p)[\mathbb{1}\otimes\Lambda](\varrho)\ , \tag{19}$$

where $p \geq d^4|\,\lambda\,|/(d^4|\,\lambda\,|+1)$, and $\lambda<0$ is the most negative eigenvalue, obtained after acting with the original map on a maximally entangled state.

We thus arrive at a modified separability criterion: ϱ is separable iff $\widetilde{[\mathbb{1}\otimes\Lambda]}(\varrho) \geq \frac{d^2\lambda}{d^4\lambda+1}$ for all positive maps Λ. An important property of this approach is that in order to check this criterion, one only has to estimate a single parameter, namely the smallest eigenvalue of the state after applying the physical map. A problem of this entanglement criterion is that it requires the application of non-local operations, which are difficult to implement. Therefore, this criterion has not yet been realised in an experiment.

2. Usefulness of Multipartite Entanglement

2.1. Multipartite Entanglement

One can define the notion of multipartite entanglement as a generalisation of the bipartite case: A pure multipartite state $|\,\psi\rangle \in \mathcal{H}_A\otimes\mathcal{H}_B\otimes ...\otimes\mathcal{H}_N$ is fully separable or n-separable, iff

$$|\,\psi\rangle \;=\; |\,a\rangle\otimes|\,b\rangle\otimes ...\otimes|\,n\rangle. \tag{20}$$

If $|\,\psi\rangle$ can be written as a tensor product of $n-1$ subsystems, it is called $n-1$-separable, and so forth. Note that k-separability implies $k-1$-separability. The state $|\,\psi\rangle$ is called biseparable iff $|\,\psi\rangle \;=\; |\,s_1\rangle\otimes|\,s_2\rangle$. If $|\,\psi\rangle$ is not biseparable , it is genuinely multipartite entangled.

A mixed state ϱ, acting on $\mathcal{H}_A\otimes\mathcal{H}_B\otimes ...\otimes\mathcal{H}_N$, is fully separable or n-separable, iff

$$\varrho \;=\; \sum_i p_i|\,a_i\rangle\langle a_i\,|\otimes|\,b_i\rangle\langle b_i\,|\otimes ...\otimes|\,n_i\rangle\langle n_i\,|, \tag{21}$$

where $0\leq p_i\leq 1$, $\sum_i p_i = 1$. Again, if ϱ can be written as a mixture of projectors onto tensor products of $n-1$ subsystems, it is called $n-1$-separable, and so forth. The state ϱ is called biseparable iff $\varrho = \sum_i p_i|\,s_{1i}\rangle\langle s_{1i}\,|\otimes|\,s_{2i}\rangle\langle s_{2i}\,|$. Here, the terms contributing to the mixture can have different splittings. If ϱ is not biseparable, it is genuinely multipartite entangled.

Unfortunately, we do not have many operational separability criteria in the multipartite case. Even for pure states, the Schmidt decomposition does not always exist. For mixed states, the Peres-Horodecki criterion of a positive partial transpose is still necessary for separability across a bipartite split.

2.2. A Zoo of Multipartite Entangled States

Various multipartite entangled states have been studied in the literature, and produced in experiments. By looking at their structure, we can understand the main properties of some of the "animals" in the multipartite entanglement zoo.

A well-known, but very vulnerable state is the GHZ state

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle), \tag{22}$$

here for three qubits. Loosing (i.e. tracing out) one of the particles means that no entanglement remains for the other subsystems. This state is as fragile as a butterfly.

A state that is inequivalent to the GHZ state is the W state,

$$|W\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle), \tag{23}$$

here for three qubits. Loosing one of the particles here means that some entanglement remains for the other subsystems. Like a rainworm, this state can be cut into two pieces, and still remain entangled.

The singlet state for four qubits reads

$$|\Psi^{(4)}\rangle = \frac{1}{\sqrt{3}}(|0011\rangle + |1100\rangle - |\psi^+\rangle|\psi^+\rangle), \tag{24}$$

where $|\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ is a Bell state. When changing the basis in all subsystems in the same way, this state remains invariant - like the chamaleon, the colour may change, but the shape remains.

Finally, an important animal in the zoo of entangled states is a cluster state, for example for four qubits:

$$|C_4\rangle = \frac{1}{\sqrt{4}}(|+0+0\rangle + |+0-1\rangle + |-1-0\rangle + |-1+1\rangle), \tag{25}$$

where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. This state and its generalisations allow to implement the one-way quantum computer, a very powerful concept in quantum information. This is the lion in the zoo of multipartite entangled states.

We do not aim at a complete zoology of entangled states in this lecture. Regarding their usefulness, with the GHZ states one can e.g. perform quantum teleportation, multi-party quantum key distribution (QKD), secret sharing, distributed dense coding, quantum imaging, and improve frequency standards. The W states are useful for multi-party QKD, secret sharing, and distributed dense coding. Singlet states can be employed for error protection (decoherence free subspace), multi-party QKD, secret sharing, and telecloning. As mentioned above, cluster states are useful for one-way quantum computing.

In order to provide a full classification of multipartite states, one should specify the quantum information processing task according to which the states are classified. A state may be useful for one task, and not useful for another: the classification scheme may thus depend on the task.

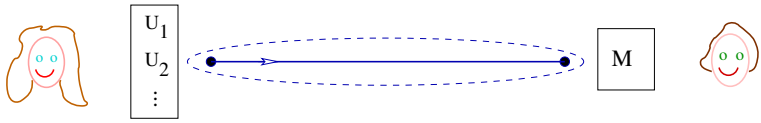


Figure 3. Original protocol for superdense coding with two parties.

2.3. Example: Distributed Dense Coding

As an example for the usefulness of multipartite entangled states for quantum information processing we will consider distributed dense coding [18].

Let us remind the reader of the original protocol for superdense coding with 2 qubits, proposed in [19]. As sketched in figure 3, we have one sender (Alice) and one receiver (Bob). Using entanglement, they manage to encode two classical bits into one qubit, sent from Alice to Bob. The protocol consists of three steps, as follows: i) Alice and Bob share the Bell state $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. ii) Alice applies one of the unitary rotations $U_i \in \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$, occurring with equal probabilities $p_i = 1/4$, to her part of the state, and sends it to Bob. iii) Bob performs a Bell measurement on the *total* state, thus gaining two bits of information.

The above protocol can be generalised to dense coding with higher dimensional bipartite states, that are not necessarily pure. Alice and Bob now share a state ρ^{AB}. Alice performs a unitary operation from the set $\{U_i\}$ with probabilities $\{p_i\}$ on her part of the system, and sends it to Bob. The ensemble of possible total states is thus $\{p_i, \rho_i^{AB}\}$ with $\rho_i^{AB} = U_i \otimes \mathbb{1}_{d_B} \rho^{AB} U_i^\dagger \otimes \mathbb{1}_{d_B}$. The Holevo bound [20] describes Bob's accessible information (which is attainable asymptotically),

$$I_{acc} = S(\overline{\rho}) - \sum_i p_i S(\rho_i^{AB}), \tag{26}$$

where $S(\xi) = -\mathrm{tr}(\xi \log \xi)$ is the von Neumann entropy of ξ, and the average state is denoted as $\overline{\rho} = \sum_i p_i \rho_i^{AB}$.

We call the maximal accessible information (optimised over all unitary encodings) the *capacity of dense coding*. For bipartite states it is given as

$$\chi = \max_{\{U_i\},\{p_i\}} I_{acc} = \log_2 d_A + S(\rho^B) - S(\rho^{AB}) \ . \tag{27}$$

This can be seen by noting that the second term in eq. (26) is equal to $-S(\rho_{AB})$, and by using the well-known fact that $S(\overline{\rho}^{AB}) \leq S(\overline{\rho}^A) + S(\overline{\rho}^B)$, together with $S(\rho^A) \leq \log_2 d_A$. The equal sign in the last two inequalities is reached by an orthogonal complete set $\{W_j\}$ with equal probabilities p_j and the trace condition $\frac{1}{d_A} \sum_j W_j^\dagger \Xi W_j = tr[\Xi]\mathbb{1}$. It is an open question whether non-unitary encodings may increase the dense coding capacity.

We can define the notion of *dense codeability* as follows. A state is useful for dense coding (dense codeable) if $\chi > \log_2 d_A$, or equivalently if

$$S(\rho^B) > S(\rho^{AB}) \ . \tag{28}$$

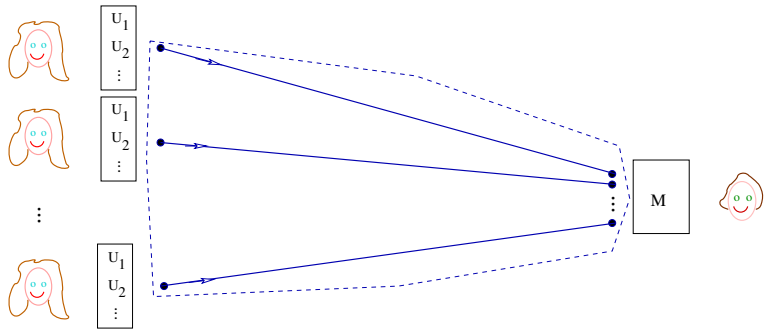


Figure 4. Distributed dense coding with N senders and one receiver.

We can make some observations and draw conclusions from the condition in eq. (28).

- For separable states, eq. (28) is never fulfilled, i.e. separable states are never useful for dense coding.
- From eq. (28) follows that ρ^{AB} violates the reduction criterion [21]. Remember that the reduction criterion reads $\mathbb{1}_{d_A} \otimes \rho^B \geq \rho^{AB}$. If the reduction criterion is violated, then the state ρ^{AB} is distillable [22]. Thus we can conclude that bound entanglement is never useful for dense coding. (For $d \times d$-systems, see [23].)
- Any *pure* bipartite entangled state is useful for dense coding.
- There exist *mixed entangled* states in dimension 2×2, which are *not* useful for dense coding, e.g. Werner states with $F < 0.748$. Thus it is not only the degree of entanglement which is important for dense codeability, but also the degree of mixedness. As quantum teleportation is possible for all entangled two-qubit states, dense coding and teleportation are inequivalent tasks.

Let us generalise dense coding to the case of N senders, so-called *distributed dense coding* [18].

The protocol for distributed dense coding is indicated in figure 4 and proceeds as follows. The N Alices and Bob share an $N+1$-partite state $\rho^{A_1 \ldots A_N B}$. Alice number j applies $U_{i_j}^{A_j}$ with probability $p_{i_j}^{A_j}$. It is important to find out whether a joint action of the senders can improve the capacity of dense coding. All Alices send their respective states to Bob, who makes a global measurement on the total state. It turns out that the Holevo bound can be achieved for *product encodings*, namely a complete orthogonal *local* set $\{W_i = \otimes_j W_{i_j}^{A_j}\}$ with the trace property as above. Thus, we arrive at the capacity of dense coding with a single receiver,

$$\chi^{A_1 \ldots A_N B} = \log_2 d_{A_1} + \ldots + \log_2 d_{A_N} + S(\rho^B) - S(\rho^{A_1 \ldots A_N B}) \; . \qquad (29)$$

Surprisingly, the senders do *not* need to perform global unitaries to achieve the capacity.

What can be said about distributed dense coding with two receivers?

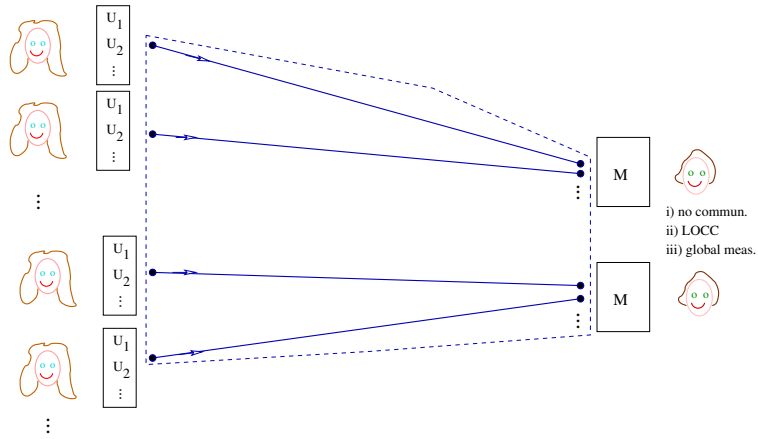


Figure 5. Distributed dense coding with N senders and two receivers.

The protocol, as sketched in figure 5, is now described as follows. The N Alices and 2 Bobs share an $N+2$-partite state $\rho^{A_1 \ldots A_N B_1 B_2}$. Alice number j applies $U_{i_j}^{A_j}$ with probability $p_{i_j}^{A_j}$ to her part of the state. The Alices $A_1, \ldots A_k$ send their states to B_1; the Alices $A_{k+1}, \ldots A_N$ send their states to B_2. There are now three different scenarios for the two Bobs, depending on which type of measurement they are allowed to make. i) No communication: the dense coding capacities are additive. ii) Local operations and classical commuication (LOCC): there is a Holevo-like upper bound, as derived in [24]. iii) Global measurement: here we are back to the situation with one receiver (with a system of higher dimension).

From these considerations we find the following bounds on the LOCC-capacity of dense coding with 2 receivers:

$$\chi^{LOCC} \geq \chi^{B_1} + \chi^{B_2} = \chi^{B_1 B_2}, \tag{30}$$

$$\chi^{LOCC} \leq \log d_{A_1} + \ldots + \log d_{A_N} + S(\rho^{B_1 B_2}) - S(\rho) = \chi^{glob}, \tag{31}$$

$$\chi^{LOCC} \leq \log d_{A_1} + \ldots + \log d_{A_N} + S(\rho^{B_1}) + S(\rho^{B_2}) - \max_{x=1,2} S(\rho^x) \tag{32}$$

$$\equiv \mathcal{B}^{LOCC} \ .$$

In eq. (32) the bound derived in [24] has been employed.

We can now introduce a classification scheme for quantum states, according to their dense codeability [18]. In figure 6 we illustrate this classification, for two receivers: The convex set "S" denotes separable states. The next shell (not including the separable states) are the bound entangled states with positive partial transpose (PBE). The shell "NBE?" are the states that have non-positive partial transpose and are conjectured to be undistillable. "D" stands for distillable states. "G-DC" are the states that are globally dense codeable. The next shell, "LOCC-DC", denotes the states that are LOCC dense codeable, but not locally dense codeable. Locally dense codeable states are denoted as "LO-DC" and form the outermost shell. Note that for one receiver the three outermost shells collaps into one shell.

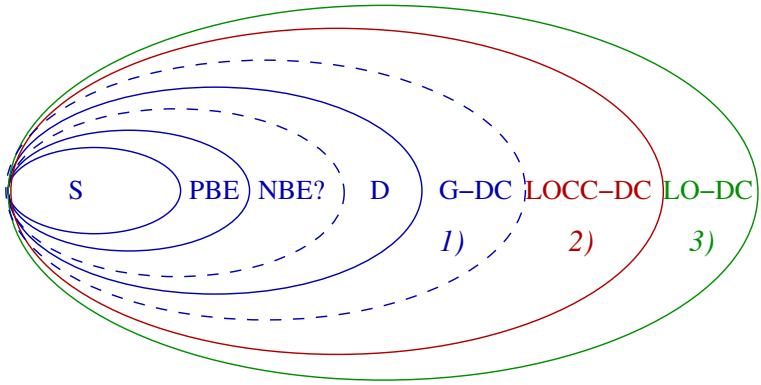


Figure 6. Classification of quantum states according to their dense codeability. For notation see main text.

We will now give examples for the three shells connected to dense coding and thus show that they are non-empty. They are also not of measure zero, because one can add a finite proportion of the identity without loosing the property of dense-codeability.

1) Example for G-DC, but not LOCC-DC:

$$|\psi_{VDDV}\rangle = \frac{1}{2}\left(|0000\rangle + |0101\rangle + |1000\rangle + |1110\rangle\right). \tag{33}$$

2) Example for LOCC-DC, but not LO-DC:

$$|\psi_{GHZ}\rangle = \tfrac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle). \tag{34}$$

The upper bound given in eq. (32) is $\chi^{LOCC} \leq 3$. This can be shown to be reachable by an explicit protocol, and therefore $\chi^{LOCC} = 3$.

3) Example for LO-DC:

$$|\psi_{2Bell}\rangle = \tfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \otimes \tfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \tag{35}$$

We have thus given a classification of multipartite quantum states according to a specific quantum information processing task. Different tasks are expected to lead to different classifications.

2.4. Detecting Multipartite Entanglement Via Witness Operators

In this subsection we connect the topics of sections 1 and 2, and address the problem of detecting multipartite entanglement.

Generalised Bell inequalities for multipartite states have been designed for example in [25,26,27] and tested experimentally. The Mermin-Klyshko inequalities are a straightforward generalisation of the CHSH inequality. Each of the n parties has two measurements with two outcomes each. We denote the dichotomic

measurements as A_i and A_i', with $i = 1, ..., n$, and the outcomes as $a_i, a_i' = \pm 1$. Then one can define a correlation operator in a recursive way,

$$M_n = \tfrac{1}{2} M_{n-1}(a_n + a_n') + \tfrac{1}{2} M_{n-1}'(a_n - a_n') \ , \tag{36}$$

with $M_1 = a_1, \ M_1' = a_1'$. Using the same line of argument as for CHSH one finds that for a local hidden variable model

$$|\langle M_n \rangle| \leq 1 \tag{37}$$

has to hold. This inequality is violated by many genuinely multipartite entangled states. For example, the GHZ-state leads to a maximal violation, namely $\langle GHZ|M_n|GHZ\rangle = 2^{(n-1)/2}$. However, there are multipartite states that do not violate the Mermin-Klyshko inequality, and thus it does not allow for general detection and/or classification of genuine multipartite entanglement.

The method of detecting multipartite entanglement via witness operators has been developed in [9,28]. The idea is very similar to the one described for the detection of bipartite entanglement in section 1.3: one first constructs a suitable witness operator $\mathcal{W}$. This operator is constructed such that it is positive on all biseparable states, and therefore detects genuine multipartite entanglement. For the construction of the witness one has to take into account the convex structure of the set of states. For three qubits this has been described in [29]. In order to find out whether a given state ϱ is multipartite entangled, one then measures the expectation value of the witness. If

$$Tr(\mathcal{W}\varrho) < 0 \ , \tag{38}$$

it follows that ϱ is genuinely multipartite entangled. In order to be able to measure the witness with present-day technology, it is decomposed locally. The *optimal* decomposition is the one that needs the minimal number of local projection measurements.

In a collaboration between the theory group of M. Lewenstein at the University of Hannover and the experimental group of H. Weinfurter at LMU Munich witnesses were implemented for the first time for the detection of multipartite entanglement [8]. The states that were experimentally created with polarised photons are the W state and the four-photon singlet state, which read

$$|\, W\rangle = (|\, 001\rangle + |\, 010\rangle + |\, 100\rangle)/\sqrt{3} \tag{39}$$

and

$$|\, \Psi^{(4)}\rangle = \frac{1}{\sqrt{3}} \left(|\, 0011\rangle + |\, 1100\rangle - \frac{1}{2}(|\, 0110\rangle + |\, 1001\rangle + |\, 0101\rangle + |\, 1010\rangle) \right) \ . \tag{40}$$

The witnesses for their detection are given as

$$\mathcal{W}_W^{(1)} = \frac{2}{3}\mathbb{1} - |\,W\rangle\langle W\,|\;, \tag{41}$$

$$\mathcal{W}_W^{(2)} = \frac{1}{2}\mathbb{1} - |\,\overline{GHZ}\rangle\langle\overline{GHZ}\,|\;, \tag{42}$$

$$\mathcal{W}_{\Psi^{(4)}} = \frac{3}{4}\mathbb{1} - |\,\Psi^{(4)}\rangle\langle\Psi^{(4)}\,|\;, \tag{43}$$

where $|\,\overline{GHZ}\rangle$ denotes a GHZ state in a rotated basis. These witnesses have been decomposed locally [28].

An example for the minimal decomposition of the witness in eq. (41) is given as

$$\begin{aligned}\mathcal{W}_W &= \frac{2}{3}\mathbb{1} - |\,W\rangle\langle W\,| \\ &= \frac{1}{24}\Big(17\cdot\mathbb{1}^{\otimes 3} + 7\cdot\sigma_z^{\otimes 3} + 3\cdot\big(\sigma_z\mathbb{1}\mathbb{1} + \mathbb{1}\sigma_z\mathbb{1} + \mathbb{1}\mathbb{1}\sigma_z\big) \\ &\quad +5\cdot\big(\sigma_z\sigma_z\mathbb{1} + \sigma_z\mathbb{1}\sigma_z + \mathbb{1}\sigma_z\sigma_z\big) \\ &\quad -(\mathbb{1}+\sigma_z+\sigma_x)^{\otimes 3} - (\mathbb{1}+\sigma_z-\sigma_x)^{\otimes 3} \\ &\quad -(\mathbb{1}+\sigma_z+\sigma_y)^{\otimes 3} - (\mathbb{1}+\sigma_z-\sigma_y)^{\otimes 3}\Big)\;,\end{aligned} \tag{44}$$

which needs only 5 measurement settings. For the interpretation as projection measurements, remember that

$$\begin{aligned}Tr[\sigma_z\otimes\sigma_z\rho] = \langle z^+z^+\,|\rho|\,z^+z^+\rangle - \langle z^+z^-\,|\rho|\,z^+z^-\rangle \\ -\langle z^-z^+\,|\rho|\,z^-z^+\rangle + \langle z^-z^-\,|\rho|\,z^-z^-\rangle\;.\end{aligned} \tag{45}$$

Note that tripartite entanglement with noise can be detected with the same witness, as long as the noise is not too large. As $Tr(\mathcal{W}_W|\,W\rangle\langle W\,|) = -1/3$, the state $\varrho = p|\,W\rangle\langle W\,| + (1-p)\mathbb{1}/8$ is detected to be genuinely multipartite entangled for all $p > 13/21$.

The experimental results when measuring the above witnesses are as follows:

$$Tr(\mathcal{W}_W^{(1)}\rho_W)_{exp} = -0.197 \pm 0.018 \quad (\text{ theory: } = -1/3)\;, \tag{46}$$

$$Tr(\mathcal{W}_W^{(2)}\rho_W)_{exp} = -0.139 \pm 0.030 \quad (\text{ theory: } = -1/4)\;, \tag{47}$$

$$Tr(\mathcal{W}_{\Psi^{(4)}}\rho_{\Psi^{(4)}})_{exp} = -0.151 \pm 0.01 \qquad (\text{ theory: } = -1/4)\;. \tag{48}$$

Thus, the existence of genuine multiparticle entanglement has been confirmed beyond any doubt. Note that the systematic shift of the experimental values with respect to the theoretical expectation is due to noise, which moves the negative expectation value of a witness towards zero. These results also confirm in a very convincing way the power of the tool of entanglement witnesses.

Acknowledgements

The research presented in this lecture resulted from collaborations with Mohamed Bourennane, Mauro D'Ariano, Manfred Eibl, Artur Ekert, Sascha Gärtner, Ot-

fried Gühne, Philipp Hyllus, Christian Kurtsiefer, Maciej Lewenstein, Chiara Macchiavello, Anna Sanpera, Aditi Sen(De), Ujjwal Sen, and Harald Weinfurter, whom I wish to thank for many enlightening and exciting discussions. This work was supported in part by the EU IST project QUPRODIS.

References

[1] R. Werner, Phys. Rev. A **40**, 4277 (1989).
[2] M. Lewenstein et al, J. Mod. Opt. **47**, 2841 (2000); D. Bruß, J. Math. Phys. **43**, 4237 (2002); D. Bruß et al, J. Mod. Opt. **49**, 1399 (2002).
[3] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).
[4] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
[5] J. Clauser, M. Horne, A. Shimony, and R. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[6] B. Terhal, Phys. Lett. A **271**, 319 (2000).
[7] M. Lewenstein, B. Kraus, I.J. Cirac, and P. Horodecki; Phys. Rev. A **62**, 052310 (2000).
[8] M. Bourennane, M. Eibl, C. Kurtsiefer, H. Weinfurter, S. Gärtner, O. Gühne, P. Hyllus, D. Bruß, M. Lewenstein, and A. Sanpera, Phys. Rev. Lett. **92**, 087902 (2004).
[9] O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello and A. Sanpera, Phys. Rev. A **66**, 062305 (2002).
[10] M. Barbieri et al, Phys. Rev. Lett. **91**, 227901 (2003).
[11] P. Hyllus, O. Gühne, D. Bruß, and M. Lewenstein, Phys. Rev. A **72**, 012321 (2005).
[12] H. Hofmann and S. Takeuchi, Phys. Rev. A **68**, 032103 (2003).
[13] O. Gühne, Phys. Rev. Lett. **92**, 117903 (2004).
[14] R. Horodecki and P. Horodecki, Phys. Lett. A **194**, 147 (1994).
[15] V. Giovanetti, Phys. Rev. A **70**, 012102 (2004).
[16] O. Gühne and M. Lewenstein, Phys. Rev. A **70**, 022316 (2004).
[17] P. Horodecki and A. Ekert, Phys. Rev. Lett. **89**, 127902 (2002).
[18] D. Bruß, G. M. D'Ariano, M. Lewenstein, C. Macchiavello, A. Sen(De), and U. Sen, Phys. Rev. Lett. **93**, 210501 (2004).
[19] C.H. Bennett and S.J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
[20] A.S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).
[21] K. Vollbrecht and M. Wolf, quant-ph/0202058.
[22] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).
[23] M. Horodecki et al, Quant. Inf. Comp. **1**, 70 (2001).
[24] P. Badziąg, M. Horodecki, A. Sen(De), and U. Sen, Phys. Rev. Lett. **91**, 117901 (2003).
[25] N. Mermin, Phys. Rev. Lett. **65**, 1838 (1990).
[26] A. Belinskii and D. Klyshko, Phys. Usp. **36**, 653 (1993).
[27] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, Phys. Rev. Lett. **88**, 170405 (2002).
[28] O. Gühne et al, J. Mod. Opt. **50**, 1079 (2003).
[29] A. Acín, D. Bruß, M. Lewenstein and A. Sanpera, Phys. Rev. Lett. **87**, 040401 (2001).

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

On Quantum Cryptography With Bipartite Bound Entangled States

Paweł Horodecki [a], Remigiusz Augusiak [a],
[a] *Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, Narutowicza 11/12, PL 80-952 Gdańsk, Poland*

Abstract. Recently the explicit applicability of bound entanglement in quantum cryptography has been shown. In this paper some of recent results respecting this topic are reviewed. In particular relevant notions and definitions are reminded. A new construction of bound entangled states containing secure correlations is presented. It provides low-dimensional $6 \otimes 6$ bound entangled states with nonzero distillable key.

Keywords. quantum cryptography, quantum security, secure key distillation, bound entanglement

1. Introduction

The explicit application of quantum entanglement in quantum information theory was the cryptographic protocol by Ekert [1]. The essential point of the protocol (cf. further modification [2]) was the entanglement monogamy principle (see [3]) which says that if the two particles are maximally entangled with each other then they are completely unentangled with any other (third) party. Hence, results of any correlation measurements on both particles must be completely safe from the cryptographic point of view, as they are uncorrelated form results of any other measurement performed on the rest of the world. This point was further exploited in a nice application [4] of entanglement distillation [5] (cf. [6]). The idea of Ref. [4] is called quantum privacy amplification (QPA). Given stationary source of pure states $\big|\Psi_{ABE}\big\rangle$ describing Alice, Bob (which are cooperating) and Eve (eavesdropper) quantum correlations protocol QPA is focused on distilling maximally entangled states

$$\big|\Psi_{+}^{(d)}\big\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} \big|i\big\rangle \otimes \big|i\big\rangle \tag{1.1}$$

from entangled quantum states

$$\varrho_{AB} = \mathrm{Tr}_E \big|\Psi_{ABE}\big\rangle\big\langle\Psi_{ABE}\big|. \tag{1.2}$$

The distillation protocol uses local operations and classical communication (LOCC) which in presence of Eve are usually called local operations and public communication (LOPC). Once Alice and Bob distill maximally correlated states $\left|\Psi_+^{(d)}\right\rangle$ by entanglement monogamy they share $\log d$ bits of classical secure bits. This can be done by performing local measurements on the state (1.1) in standard bases $\{|i\rangle_A\}_{i=0}^{d-1}$ and $\{|i\rangle_B\}_{i=0}^{d-1}$.

However, since 1998 it has been known that there is quantum entanglement called bound entanglement that can not be distilled to a pure form [7]. For a long time there was a common belief that distillation of secure key from quantum state is possible only when QPA is. In other words, that bound entanglement is useless for quantum cryptography. In fact the results of extensive analysis of two qubit case [8] naturally suggested equivalence of entanglement distillation protocols and secure key distillation.

Surprisingly it is not true, as it has been shown in papers [9,10]. Before we shall recall main observations of the latter, let us point out the key ingredient of their reasoning. Namely, why QPA might be not necessary for distilling secure key? In fact, if Alice and Bob share maximally entangled state (1.1) they in a sense have much stronger security than they need. In fact they will get secure correlations if they measure the state in *any* pair of bases of the form: $\{\left|e_i^{(A)}\right\rangle = U|i\rangle_A\}_{i=0}^{d-1}$ and $\{\left|e_i^{(B)}\right\rangle = U^*|i\rangle_B\}_{i=0}^{d-1}$, where U stands for arbitrary unitary operation. It is crucial to understand that in quantum cryptography it would be enough to have *a single basis*, measurement in which could give secure correlations. To be more specific, the security requirement is to get (via local measurement of Alice and Bob in one basis) the state that is product with Eve's degrees of freedom.

This leads us to the notion of ccq state [12], i.e., tripartite state of Alice, Bob and Eve that is, after local von Neumann measurements, classical on Alice and Bob parts and quantum on Eve part. With this notion one can summarize the idea of [9,10] as follows: Alice and Bob should proceed to distill such a state $\gamma_{ABA'B'}^{(2)}$ that, after considering its purification $\left|\Psi_{ABA'B'E}\right\rangle$ and performing a local measurements in a standard product basis on its AB part the resulting ccq state ϱ_{ABE} is (i) product with respect to the division $AB|E$ (ii) contains maximal classical correlations between Alice and Bob.

In this work we shall briefly describe main mathematical elements of the construction [9,10], recall the idea of one-way distillation provided in [11] and provide a new construction of bound entangled states with secure quantum key.

2. Main Notions of General Secure Key Distillation Schemes

Here we shall remind and discuss main notions introduced in Refs. [9,10].

Assume that Alice and Bob wish to communicate but without participation of an eavesdropper Eve. To this aim, as a source of quantum correlations, they use a quantum state $\varrho_{ABA'B'}$. Here, subsystem AA' (BB') belongs to Alice (Bob). Moreover, following [9], subsystem AB ($A'B'$) shall be called the *key part (shield part)*. To make the considerations more formal, each subsystem of $\varrho_{ABA'B'}$ shall be represented by respective Hilbert space, i.e., Alice's subsystems by $\mathscr{H}_A$ and $\mathscr{H}_{A'}$ and Bob's by $\mathscr{H}_B$ and $\mathscr{H}_{B'}$. Hence, the key part of $\varrho_{ABA'B'}$ is defined on

$\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$ and the shield part on $\mathscr{H}' = \mathscr{H}_{A'} \otimes \mathscr{H}_{B'}$. Hereafter we shall be assuming that $\dim \mathscr{H}_A = \dim \mathscr{H}_B = d$, $\dim \mathscr{H}_{A'} = d_{A'}$, and $\dim \mathscr{H}_{B'} = d_{B'}$.

Firstly, let us remind the notion of ccq state. To this aim let us introduce a product basis defined on Hilbert space $\mathscr{H}_A \otimes \mathscr{H}_B \sim \mathbb{C}^d \otimes \mathbb{C}^d$, i.e.,

$$\mathscr{B}_{AB} = \left\{ \left|e_i^{(A)}\right\rangle \otimes \left|e_j^{(B)}\right\rangle \right\}_{i,j=0,\ldots,d-1}, \tag{2.3}$$

where $\{\left|e_i^{(A)}\right\rangle\}_{i=0,\ldots,d-1}$ and $\{\left|e_i^{(B)}\right\rangle\}_{i=0,\ldots,d-1}$ are arbitrary bases spanning the Hilbert spaces, respectively, $\mathscr{H}_A$ and $\mathscr{H}_B$. Of course, these bases may be chosen to be standard and therefore $\mathscr{B}_{AB}$ is the standard basis in $\mathbb{C}^d \otimes \mathbb{C}^d$. Then we have the following

Definition 1. *(ccq state) We call the state $\tilde{\varrho}_{ABE}$ a* **ccq state** *of initial state $\varrho_{ABA'B'}$ with respect to the basis $\mathscr{B}_{AB}$ if $\tilde{\varrho}_{ABE}$ is a result of measurement of $\varrho_{ABE} = \mathrm{Tr}_{A'B'}\left|\Psi_{ABA'B'E}\right\rangle\left\langle\Psi_{ABA'B'E}\right|$ in the product basis $\mathscr{B}_{AB}$. Here $\left|\Psi_{ABA'B'E}\right\rangle$ is a purification of $\varrho_{ABA'B'}$.*

As an illustrative example let us consider a density matrix $\eta_{ABA'B'}$ acting on $(\mathbb{C}^2)^{\otimes 4}$ and given by $\eta_{ABA'B'} = p\left|0111\right\rangle\left\langle 0111\right| + (1-p)\left|1000\right\rangle\left\langle 1000\right|$. As one may easily verify its standard purification takes the form $\left|\Psi_{ABA'B'E}\right\rangle = \sqrt{p}\left|01110\right\rangle + \sqrt{1-p}\left|10001\right\rangle$. Therefore, the ccq state of $\eta_{ABA'B'}$ with respect to standard basis $\mathscr{B}_{AB}^{(\mathrm{st})} \equiv \{|i\rangle_A|j\rangle_B\}_{i,j=0}^{1}$ is $\tilde{\eta}_{ABE} = p\left|010\right\rangle\left\langle 010\right| + (1-p)\left|101\right\rangle\left\langle 101\right|$.

Now, one can ask when a given state is said to be secure. As an answer one gives the following (see [10])

Definition 2. *(security) We call the state $\varrho_{ABA'B'}$* **secure** *with respect to the basis $\mathscr{B}_{AB}$ if its ccq state is of the form*

$$\tilde{\varrho}_{ABE} = \left[\sum_{i,j=0}^{d-1} p_{ij} \left|e_i^{(A)} e_j^{(B)}\right\rangle\left\langle e_i^{(A)} e_j^{(B)}\right| \right] \otimes \varrho_E. \tag{2.4}$$

The security of such a state follows from the fact that Eve is completely uncorrelated from distribution represented by AB system after Alice and Bob measurement. Note that if the distribution p_{ij} is homogenous, i.e., $p_{ij} = 1/d; (i,j = 0,\ldots,d-1)$, then we say that $\varrho_{ABA'B'}$ has a $\mathscr{B}_{AB}$**-key**.

A very important ingredient of construction discussed here is a special class of controlled unitary operations (see [9,10]) that we recall by the following

Definition 3. *(twisting) Let $\mathsf{U}_{ij}^{(A'B')}$ be certain unitary operations acting on subsystem $A'B'$. For a given basis $\mathscr{B}_{AB}$ we call the operation*

$$\mathsf{U} = \sum_{i,j=0}^{d-1} \left|e_i^{(A)} e_j^{(B)}\right\rangle\left\langle e_i^{(A)} e_j^{(B)}\right| \otimes \mathsf{U}_{ij}^{(A'B')} \tag{2.5}$$

$\mathscr{B}_{AB}$**-twisting** *or shortly,* **twisting**.

The importance of such a class of operations follows from the fact that applied to a given state $\varrho_{ABA'B'}$, U *do not* change its ccq state. It means that if we take $\varrho_{ABA'B'}$ and $\sigma_{ABA'B'} = \mathsf{U}\varrho_{ABA'B'}\mathsf{U}^\dagger$, their ccq states are exactly the same, i.e., $\tilde{\varrho}_{ABE} = \tilde{\sigma}_{ABE}$.

Now we define the central notion of the generalised approach provided in [9,10]. This is the notion of private state that has $\log d$ bits of secure key encoded in its AB part of $d \otimes d$ type).

Definition 4. *(pdit) Let $\varrho_{ABA'B'}$ is a density operator acting on the Hilbert space $\mathscr{H} \otimes \mathscr{H}'$ and $\varrho_{A'B'}$ is a density matrix acting on $\mathscr{H}'$. Let $\mathsf{U}_i^{(A'B')}$ $(i = 0, \ldots, d-1)$ be certain unitary operations acting on $A'B'$ system. Then we call the state $\varrho_{ABA'B'}$* **private state** *or* **pdit** *with respect to the basis $\mathscr{B}_{AB}$ if it is of the form*

$$\varrho_{ABA'B'} = \frac{1}{d}\sum_{i,j=0}^{d-1} \left|e_i^{(A)}e_i^{(B)}\right\rangle\!\left\langle e_j^{(A)}e_j^{(B)}\right| \otimes \mathsf{U}_i^{(A'B')}\varrho_{A'B'}\mathsf{U}_j^{(A'B')\dagger}. \tag{2.6}$$

Hereafter, as usual, we shall denote the private dits by $\gamma^{(d)}_{ABA'B'}$. In the case when the dimension of the key part is $d = 2$ on each side, we have to do with **private bit** or **pbit**.

Now it is important to pose the question: *What is the essential feature that allows private bit to be truly private?* Namely, a detailed analysis shows [9,10] that ccq state of private bit $\gamma^{(d)}_{ABA'B'}$ is the same as the ccq state of the following state (called **basic pdit**)

$$\mathcal{P}_+^{(d)} \otimes \sigma_{A'B'}, \tag{2.7}$$

where $\sigma_{A'B'} = \mathrm{Tr}_{AB}\gamma^{(d)}_{ABA'B'}$ and $\mathcal{P}_+^{(d)}$ is a projector onto $\left|\Psi_+^{(d)}\right\rangle$. This is because it is always possible for a given pbit to find such a twisting under which $\gamma^{(2)}_{ABA'B'}$ transforms it to (2.7) and conversely. Thus, after performing measurement in the basis $\mathscr{B}_{AB}$ the physical system ABE is in the same state irrespective of whether before Alice and Bob state was (2.7) or just a pdit $\gamma^{(d)}_{ABA'B'}$. Hence the security with respect to the measurement in *that particular basis* is the same as if Alice and Bob really shared maximal entangled state $\mathcal{P}_+^{(d)}$! This is the key observation for understanding the essence of private dit.

We conclude the preliminary section recalling the definition of distillable key [9,10] and related theorem.

Definition 5. *(distillable key) Let σ_{AB} be a density matrix acting on $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ and let $\mathscr{P}_n$ be a sequence of LOCC operations such that $\mathscr{P}_n(\sigma_{AB}^{\otimes n}) = \Sigma_n$, where Σ_n is defined on $\mathbb{C}^{d_n} \otimes \mathbb{C}^{d_n} \otimes \mathscr{H}_{A'} \otimes \mathscr{H}_{B'}$. The set $\mathscr{P} = \{\mathscr{P}_n\}_{n=1}^{\infty}$ is said to be a* **pdit distillation protocol** *of a given state σ_{AB} if the following relation*

$$\lim_{n\to\infty} \left|\left|\Sigma_n - \gamma^{(d_n)}_{ABA'B'}\right|\right|_{\mathrm{Tr}} = 0. \tag{2.8}$$

The rate of this protocol is defined as

$$\mathcal{R}(\mathscr{P}) = \limsup_{n, d_n \to \infty} \frac{\log d_n}{n} \tag{2.9}$$

and distillable key of σ_{AB} *as*

$$K_D(\sigma_{AB}) = \sup_{\mathscr{P}} \mathcal{R}(\mathscr{P}). \tag{2.10}$$

There is a problem however, since the above definition has a complicated form. It is hard to see whether given state fulfills the above condition or not. Fortunately, one can simplify the task providing necessary and sufficient conditions for nonzero distillable key that are more operational then the definition itself. Here we provide a summary of the conditions proven in [9,10], which are enough to analyse cryptographic usefulness of many quantum states:

Theorem 1. *The following three conditions are equivalent:*

(i) finite number of copies of state $\varrho_{ABA'B'}$ *can be transformed with some LOCC protocol into the state* $\sigma_{ABA'B'}$ *arbitrarily close in the trace norm to certain pbit* $\gamma^{(2)}_{ABA'B'}$

(ii) finite number of copies of state $\varrho_{ABA'B'}$ *can be transformed with some LOCC protocol into the state* $\sigma_{ABA'B'}$ *with trace norm of the element* $\mathsf{A}^{(A'B')}_{00,11}$ *arbitrarily close to* $1/2$*, where the element is defined by the representation:*

$$\varrho_{ABA'B'} = \sum_{i,j,k,l=0}^{1} \left|e_i^{(A)} e_j^{(B)}\right\rangle\!\left\langle e_k^{(A)} e_l^{(B)}\right| \otimes \mathsf{A}^{(A'B')}_{ij,kl}. \tag{2.11}$$

(iii) the state ϱ *has nonzero distillable key, i.e., one has* $K_D(\varrho) > 0$.

Moreover any convergence of $||\mathsf{A}^{(A'B')}_{00,11}||_{\mathrm{Tr}}$ *to* $1/2$ *from (ii) during a given protocol is equivalent to a convergence of the state to a certain pbit during that protocol.*

Especially the condition (ii) serves as a useful criterion which allows to adjudge an applicability of a given state to quantum cryptography. In next section we shall illustrate its power.

3. New Class of Bound Entangled States With Secure Quantum Key

In this section we present the main results. We provide a construction of a state that is useful for quantum cryptography simultaneously being bound entangled. Note that the construction, however based on that presented in [11], is different in details from known so far [9,10,11] and sheds some light on the still unexplored domain of bound entanglement. Hereafter it is assumed that $d = 2$, i.e., the key part of $\varrho_{ABA'B'}$ consists of qubits and that $\dim \mathscr{H}_{A'} = \dim \mathscr{H}_{B'} = D$, which allows us to write $\mathscr{H}' = \mathbb{C}^D \otimes \mathbb{C}^D$. We may also assume for simplicity that $\mathscr{B}_{AB}$ is standard basis in $\mathbb{C}^2 \otimes \mathbb{C}^2$.

3.1. The Construction

To obtain a better insight into the construction we begin our considerations from an illustrative example with the shield part of dimension $D = 3$ on each side. This with the assumption that $d = 2$ makes the considered state to be of dimension 6 in each side. Finally, we show that the construction is possible for arbitrary $D \geq 3$.

3.1.1. The $6 \otimes 6$ Case

At the very beginning suppose that Alice and Bob possess the following state (cf. [11])

$$\varrho_{ABA'B'} = \frac{11}{40}\begin{bmatrix} |X_3| & 0 & 0 & X_3 \\ 0 & \left|X_3^{T_{B'}}\right|^{T_{B'}} & 0 & 0 \\ 0 & 0 & \left|X_3^{T_{B'}}\right|^{T_{B'}} & 0 \\ X_3 & 0 & 0 & |X_3| \end{bmatrix} \tag{3.12}$$

where the superscript $T_{B'}$ denotes the partial transposition with respect to the system B' and X_3 is a symmetric 9×9 matrix of the form

$$X_3 = \frac{1}{11}\begin{bmatrix} -1&0&0&0&1&0&0&0&1 \\ 0&1&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0 \\ 1&0&0&0&-1&0&0&0&1 \\ 0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&1&0 \\ 1&0&0&0&1&0&0&0&-1 \end{bmatrix}. \tag{3.13}$$

which is defined on the Hilbert space $\mathscr{H}_{A'} \otimes \mathscr{H}_{B'} = \mathbb{C}^3 \otimes \mathbb{C}^3$. In order to show that the matrix $\varrho_{ABA'B'}$ is Hermitian and nonnegative, i.e., represents a quantum state, let us observe that X may be decomposed as $X_3 = (1/11)(\mathcal{P}_+^{(3)} - 2\mathcal{P}^{(3)} + \mathcal{Q}^{(3)})$. Here $\mathcal{P}_+^{(3)}$ is a projector onto the maximally entangled state $\left|\Psi_+^{(3)}\right\rangle = (1/\sqrt{3})\sum_{i=0}^{2}\left|i\right\rangle\left|i\right\rangle$ belonging to $\mathbb{C}^3 \otimes \mathbb{C}^3$ and $\mathcal{P}^{(3)}$, and $\mathcal{Q}^{(3)}$ are projectors given by relations

$$\mathcal{P}^{(3)} = \sum_{i=0}^{2}\left|ii\right\rangle\left\langle ii\right| - \mathcal{P}_+^{(3)}, \qquad \mathcal{Q}^{(3)} = \mathcal{I}_9 - \sum_{i=0}^{2}\left|ii\right\rangle\left\langle ii\right|, \tag{3.14}$$

where $\mathcal{I}_9$ stands for an identity acting on the Hilbert space $\mathbb{C}^3 \otimes \mathbb{C}^3$. These above projectors are orthogonal and therefore one obtains

$$|X_3| = \frac{1}{11}\left[\mathcal{P}_+^{(3)} + 2\mathcal{P}^{(3)} + \mathcal{Q}^{(3)}\right] = \frac{1}{11}\begin{bmatrix} \frac{5}{3} & 0\,0\,0 & -\frac{1}{3} & 0\,0\,0 & -\frac{1}{3} \\ 0 & 1\,0\,0 & 0 & 0\,0\,0 & 0 \\ 0 & 0\,1\,0 & 0 & 0\,0\,0 & 0 \\ 0 & 0\,0\,1 & 0 & 0\,0\,0 & 0 \\ -\frac{1}{3} & 0\,0\,0 & \frac{5}{3} & 0\,0\,0 & -\frac{1}{3} \\ 0 & 0\,0\,0 & 0 & 1\,0\,0 & 0 \\ 0 & 0\,0\,0 & 0 & 0\,1\,0 & 0 \\ 0 & 0\,0\,0 & 0 & 0\,0\,1 & 0 \\ -\frac{1}{3} & 0\,0\,0 & -\frac{1}{3} & 0\,0\,0 & \frac{5}{3} \end{bmatrix}. \tag{3.15}$$

From the above equation one infers two facts, first that the trace norm[1] of X_3 is $||X_3||_{\mathrm{Tr}} = 1$ and the second that $|X_3|^{T_{B'}} \geq 0$. Moreover, the matrix X_3 partially transposed with respect to the subsystem B' may be written in the form $X_3^{T_{B'}} = (1/11)\left[2\mathcal{S}^{(3)} - \left(\mathcal{I}_9 - \mathcal{Q}^{(3)}\right)\right]$, where $\mathcal{S}^{(3)}$ is a projector given by

$$\mathcal{S}^{(3)} = \frac{1}{2}\left[\mathcal{I}_9 + \mathcal{V}^{(3)} - 2\sum_{i=0}^{2}|ii\rangle\langle ii|\right] \tag{3.16}$$

with $\mathcal{V}^{(3)}$ being the swap operator defined by relation $\mathcal{V}^{(3)}|\varphi_1\rangle|\varphi_2\rangle = |\varphi_2\rangle|\varphi_1\rangle$ for $|\varphi_i\rangle \in \mathbb{C}^3$ $(i = 1, 2)$. Note that $\mathcal{I}_9 - \mathcal{Q}^{(3)}$ and $\mathcal{S}^{(3)}$ are orthogonal projectors and therefore $\left|X_3^{T_{B'}}\right| = (1/11)\left[2\mathcal{S}^{(3)} + \mathcal{I}_9 - \mathcal{Q}^{(3)}\right]$. Since $\left|X_3^{T_{B'}}\right|^{T_{B'}} = (1/11)\left[3\mathcal{P}_+^{(3)} + \mathcal{Q}^{(3)}\right]$, one may easily conclude that $\left|X_3^{T_{B'}}\right|^{T_{B'}} \geq 0$. This with the aid of the fact that the matrix X_3 is symmetric and real, ensures that $\varrho_{ABA'B'}$ represents a quantum state.

Now we are in position to prove that $\varrho_{ABA'B'}$ satisfies PPT (positive partial transpose) criterion [13,14]. To this aim we show that transposition with respect to subsystem BB' preserves the positivity. Indeed, the state $\varrho_{ABA'B'}$ transposed with respect to the BB' subsystem, remains a positive operator. To see this fact explicitly, we write

$$\varrho_{ABA'B'}^{T_{BB'}} = \frac{11}{40}\begin{bmatrix} |X_3|^{T_{B'}} & 0 & 0 & 0 \\ 0 & \left|X_3^{T_{B'}}\right| & X_3^{T_{B'}} & 0 \\ 0 & X_3^{T_{B'}} & \left|X_3^{T_{B'}}\right| & 0 \\ 0 & 0 & 0 & |X_3|^{T_{B'}} \end{bmatrix}. \tag{3.17}$$

Positivity of the above operator stems from two facts. As previously mentioned $|X_3|^{T_{B'}} \geq 0$ and on the other hand the off-diagonal elements are blocked by $\left|X_3^{T_{B'}}\right|$.

3.1.2. Construction of General $2D \otimes 2D$ Case

Trying to generalize the above investigations we start from the symmetric matrix

[1] For an arbitrary matrix Ξ the trace norm is defined by relation $||\Xi||_{\mathrm{Tr}} = \mathrm{Tr}\sqrt{\Xi^\dagger\Xi}$.

$$X_D = \frac{1}{D^2+2D-4}\left[(D-2)\mathcal{P}_+^{(D)} - 2\mathcal{P}^{(D)} + \mathcal{Q}^{(D)}\right], \tag{3.18}$$

where, as previously, $\mathcal{P}_+^{(D)}$ represents a projector onto maximally entangled state $\left|\Psi_+^{(D)}\right\rangle$. Orthogonal projectors $\mathcal{P}^{(D)}$ and $\mathcal{Q}^{(D)}$ are given by

$$\mathcal{P}^{(D)} = \sum_{i=0}^{D-1} \left|ii\right\rangle\!\left\langle ii\right| - \mathcal{P}_+^{(D)}, \qquad \mathcal{Q}^{(D)} = \mathcal{I}_{D^2} - \sum_{i=0}^{D-1} \left|ii\right\rangle\!\left\langle ii\right|. \tag{3.19}$$

From Eq. (3.18) we have

$$|X_D| = \frac{1}{D^2+2D-4}\left[(D-2)\mathcal{P}_+^{(D)} + 2\mathcal{P}^{(D)} + \mathcal{Q}^{(D)}\right] \tag{3.20}$$

and therefore $|X_D|^{T_{B'}} \geq 0$. Subsequently, after elementary steps, we may obtain $X_D^{T_{B'}} = \left[1/(D^2+2D-4)\right]\left[2\mathcal{S}^{(D)} - \left(\mathcal{I}_{D^2} - \mathcal{Q}^{(D)}\right)\right]$ with $\mathcal{S}^{(D)}$ being defined as

$$\mathcal{S}^{(D)} = \frac{1}{2}\left[\mathcal{I}_{D^2} + \mathcal{V}^{(D)} - 2\sum_{i=0}^{D-1} \left|ii\right\rangle\!\left\langle ii\right|\right], \tag{3.21}$$

where $\mathcal{V}^{(D)}$ is the swap operator acting on $\mathbb{C}^D \otimes \mathbb{C}^D$. Again, the projectors $\mathcal{I}_{D^2} - \mathcal{Q}^{(D)}$ and $\mathcal{S}^{(D)}$ are orthogonal and therefore

$$\left|X_D^{T_{B'}}\right| = \frac{1}{D^2+2D-4}\left[2\mathcal{S}^{(D)} + \mathcal{I}_{D^2} - \mathcal{Q}^{(D)}\right]. \tag{3.22}$$

Finally, performing partial transposition with respect to subsystem B', we have $\left|X_D^{T_{B'}}\right|^{T_{B'}} = \left[1/(D^2+2D-4)\right]\left[D\mathcal{P}_+^{(D)} + \mathcal{Q}^{(D)}\right]$ and therefore $\left|X_D^{T_{B'}}\right|^{T_{B'}} \geq 0$. Now we can introduce a class of mixed states

$$\varrho_{ABA'B'}^{(D)} = \frac{1}{4}\frac{D^2+2D-4}{D^2+D-2}\begin{bmatrix} |X_D| & 0 & 0 & X_D \\ 0 & \left|X_D^{T_{B'}}\right|^{T_{B'}} & 0 & 0 \\ 0 & 0 & \left|X_D^{T_{B'}}\right|^{T_{B'}} & 0 \\ X_D & 0 & 0 & |X_D| \end{bmatrix}. \tag{3.23}$$

Again, by virtue of the fact that $|X_D|^{T_{B'}} \geq 0$ one may infer that partial transposition with respect to subsystem BB' preserves the positivity of $\varrho_{ABA'B'}^{(D)}$. In other words, the following matrix

$$\varrho_{ABA'B'}^{(D)T_{BB'}} = \frac{1}{4}\frac{D^2+2D-4}{D^2+D-2}\begin{bmatrix} |X_D|^{T_{B'}} & 0 & 0 & 0 \\ 0 & \left|X_D^{T_{B'}}\right| & X_D^{T_{B'}} & 0 \\ 0 & X_D^{T_{B'}} & \left|X_D^{T_{B'}}\right| & 0 \\ 0 & 0 & 0 & |X_D|^{T_{B'}} \end{bmatrix}. \tag{3.24}$$

possess the nonnegative eigenvalues.

3.2. Proof of Nonzero Distillable Key K_D

Since the state $\varrho^{(D)}_{ABA'B'}$ given by Eq. (3.23) is PPT one might expect that it is separable - it satisfies necessary (PPT) condition for separability [13]. However it is not. This is in agreement with the fact [14] that PPT condition is sufficient for separability only for the cases $M \otimes N$, $MN \leq 6$ while here we have $MN = (2D)^2 \geq 36$. Like in [9] we shall prove nonseparability of $\varrho^{(D)}_{ABA'B'}$ in a very nonstandard way. We simply show that the state has nonzero K_D. Such a state must be entangled since, due to seminal result of Ref. [15], any separable state has $K_D = 0$. Because state is PPT and entangled it must be bound entangled [7].

To prove the cryptographic use of $\varrho_{ABA'B'}$, below we show that there exist a LOCC protocol that allows Alice and Bob to approach arbitrarily closely to some pbit $\gamma^{(2)}_{ABA'B'}$. Note that, obviously, since LOCC operations preserves PPT property (see [7]), the resulting state is still bound entangled. Given k copies of the state $\varrho_{ABA'B'}$ in the i-th step of the protocol Alice and Bob perform the following operations:

- They take the state $\varrho^{(i-1)}_{ABA'B'}$ $(i = 1, \ldots, k-1)$ and one of remaining $k-i+1$ copies of $\varrho_{ABA'B'}$ (here $\varrho^{(0)}_{ABA'B'} = \varrho_{ABA'B'}$).
- They perform C-NOT operation treating qubits A and B of $\varrho_{ABA'B'}$ as source qubits and that of $\varrho^{(i-1)}_{ABA'B'}$ as target qubits.
- They perform a measurement of target qubits in computational basis and then compare their results. If both of them have the same results (00 or 11) then they keep the source state. Otherwise they get rid of it.

After performing all k steps, with some probability they arrive at the following state

$$\varrho^{(D,k)}_{ABA'B'} = \frac{1}{N_{D,k}} \begin{bmatrix} |X_D|^{\otimes k} & 0 & 0 & X_D^{\otimes k} \\ 0 & \left(\left|X_D^{T_{B'}}\right|^{T_{B'}}\right)^{\otimes k} & 0 & 0 \\ 0 & 0 & \left(\left|X_D^{T_{B'}}\right|^{T_{B'}}\right)^{\otimes k} & 0 \\ X_D^{\otimes k} & 0 & 0 & |X_D|^{\otimes k} \end{bmatrix}, \quad (3.25)$$

where (for $D \geq 3$)

$$N_{D,k} = 2\mathrm{Tr}|X_D|^{\otimes k} + 2\mathrm{Tr}\left(\left|X_D^{T_{B'}}\right|^{T_{B'}}\right)^{\otimes k} = 2\left[1 + \left(\frac{D^2}{D^2 + 2D - 4}\right)^k\right] \xrightarrow{k\to\infty} 2. \quad (3.26)$$

If we define, according to (2.11) the matrix, $\mathsf{A}^{(A'B')}_{00,11}(D,k) = (1/N_{D,k})X_D^{\otimes k}$ we can see that $||\mathsf{A}^{(A'B')}_{00,11}(D,k)||_{\mathrm{Tr}} \to 1/2$ whenever $k \to \infty$, for $D \geq 3$. This means that repeating the whole procedure described above, one may find such a k that the trace norm of the upper right block of $\varrho^{(D,k)}_{ABA'B'}$ is close to $1/2$ with arbitrary precision. According to the the Theorem 1 (see Section 2) this convergence guarantees that the original state $\varrho^{(D)}_{ABA'B'}$ defined by the formula (3.23) satisfies

$K_D > 0$. The Theorem 2 guarantees in particular, that (like it was in [9]) the above sequence of bound entangled states $\varrho^{(D,k)}_{ABA'B'}$ approaches private bit.

4. Summary and Discussion

We have summarized main elements of general scheme of distillation of secure key [9,10]. The central notion of the scheme is the idea of private bit (with its natural generalization - private dit) which is the state that, in general, consists of two parts: the key part AB and the shield part $A'B'$. The first contains a bit of secure key. The role of the second part is - in a sense - to protect the key form Eve. A surprising fact, found in [9] is that PPT bound entangled states can approach private bit in trace norm. Since this convergence is a necessary and sufficient condition to distill secure key form the original state, this means that bound entangled states can serve as a source of distillable key [9,10].

First bound entangled states with nonzero distillable key K_D were provided in [9]. They required however very high dimensions. The small $(4 \otimes 4)$ bound entangled states with nonzero distillable key were provided later in paper [11].

Here we have provided new class of small (of, among others, $6 \otimes 6$ type) bound entangled states with that property. We have proven this fact applying an easy criterion form [9] showing that a given sequence of quantum states approaches the sequence of private bits. The LOCC protocol applied to produce such a sequence was of two-way type. As noticed in [9] at some point the elements of the sequence got one-way distillable key which can be distilled with help of Devetak-Winter protocol. The original problem was that such states were of very high dimensions. Quite surprisingly the low-dimensional bound entangled states provided in [11] represent one-way distillable key. It was proven with help of the observation that any biased mixture of two private bits with second of them having the key part rotated locally with σ_x Pauli matrix contains nonzero distillable key. Namely one has the following theorem [11]:

Theorem 2. *For two private bits* $\gamma_1^{(2)}$, $\gamma_2^{(2)}$ *one-way distillable key of the mixture of the form*:

$$\rho = p_1\gamma_1^{(2)} + p_2\sigma_x^{(A)}\gamma_2^{(2)}\sigma_x^{(A)} \tag{4.27}$$

with, $\sigma_x^{(A)} = [\sigma_x]_A \otimes I_{A'BB'}$ *satisfies* $K_{\overrightarrow{D}}(\rho) \geq 1 - h(p_1)$ *with binary entropy* $h(p_1)$[2].

Here $K_{\overrightarrow{D}}(\rho)$ stands for cryptographic key distillable from ρ with help of forward classical communication. The natural question is whether one can modify the $6 \otimes 6$ bound entangled states provided in the present paper to get mixture of two private bits of the above form. Our first analysis has shown that most probably it is impossible to turn our example into a state of the form (4.27) while keeping bound entanglement property. In Ref. [11] the construction leading to *bound entangled* state of the form (4.27) was based on some properties of states that

[2]Binary entropy of the distribution $\{p_1, p_2\}$ is defined as $h(p_1) = -p_1 log_2 p_1 - p_2 log_2 p_2$

were used in locking entanglement measures effects. May be that was the reason why the construction was successful there. Still there is a natural question about other methods to construct low-dimensional bound entangled states with one-way distillable key.

Of course the most important open problem is whether any entangled bipartite state contains nonzero distillable key or not. In multipartite case it is not true - there are states that are entangled but no secure key between any of the parties can be distilled [16]. In bipartite case lack of such states is guaranteed (via entanglement distillation approach [4]) only for $2 \otimes 2$ and $2 \otimes 3$ cases, since, as shown in [17,18], all the entangled states can be distilled to singlet form in those cases. For $d \otimes d$ with $d \geq 4$ it is known that at least some bound entangled (ie. nondistillable to singlets) states have nonzero distillable key [11]. No example of $3 \otimes 3$ or $2 \otimes 4$ states with that property is still known.

Acknowledgements

The authors thank Maciej Demianowicz and Michał Horodecki for fruitful discussions. The work is partially supported by Polish Ministry of Scientific Research and Information Technology grant under the (solicited) project no. PBZ-MIN-008/P03/2003 and by EC grant: RESQ, contract no. IST-2001-37559.

References

[1] A. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[2] C. H. Bennett, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).
[3] B. M. Terhal, IBM J. Res. Dev. **48**, 71 (2004).
[4] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Phys. Rev. Lett. **77**, 2818 (1996).
[5] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
[6] C. H. Bennett, D. P. Di Vincenzo, J. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3814 (1997).
[7] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
[8] N. Gisin and M. Wolf, Phys. Rev. Lett. **83**, 4200 (1999); A. Acin, L. Massanes, and N. Gisin, quant-ph/0303053.
[9] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **94**, 160502 (2005).
[10] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, quant-ph/0506189.
[11] K. Horodecki, Ł. Pankowski, M. Horodecki, and P. Horodecki, quant-ph/0506203.
[12] I. Devetak and A. Winter, Proc. R. Soc. A **461**, 207 (2005).
[13] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).
[14] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
[15] M. Curty, M. Lewenstein, and N. Lütkenhaus, Phys. Rev. Lett. **92**, 217903 (2004).
[16] R. Augusiak and P. Horodecki, quant-ph/0405187; R. Augusiak and P. Horodecki, Phys. Rev. A **73**, 012318 (2005).
[17] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **78**, 574 (1997).
[18] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4026 (1999).

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Unitary Local Permutations on Bell Diagonal States of Qudits and Quantum Distillation Protocols

Hector Bombin, Miguel A. Martin-Delgado
Departamento de Física Teórica I, Universidad Complutense, 28040. Madrid, Spain

Abstract. Entanglement distillation protocols are generalized for qudits of arbitrary dimension by studying the group of unitary local permutations. We introduce the concept of *joint performance parameter* η that allows us the comparison of distillation protocols with different values of fidelity, probability of success and number of Bell pairs used altogether.We analyze several distillation protocols assisted with twirling operations as the dimension D of qudits vary. We find that the best performance according to η is not achieved for qubits ($D = 2$), but for qutrits ($D = 3$) and $n = 3$ input pairs of Bell states. We propose and study an extension of the Quantum Privacy Amplification protocols that work for arbitrary dimension D, being more efficient when D is a prime number.

Keywords. Entanglement, Distillation, Qudits, Quantum Privacy Amplification

1. Introduction

We show that the analysis of entanglement distillation protocols [1], [2] for qudits of arbitrary dimension D benefits from applying basic concepts from number theory [3]. This is so because the set modulus $\mathbf{Z}_D := \mathbf{Z}/D\mathbf{Z}$ can be identified with the elements of a basis of $\mathbf{C}^D$. We characterize the group of local permutations over tensor products of Bell states as an affine symplectic matrix group. Then we find that a partition of $\mathbf{Z}_D^n$ into divisor classes is related to the invariant properties of mixed Bell diagonal states under this group. We construct a very general class of recursion protocols by means of unitary operations implementing these local permutations. We analyze these distillation protocols depending on whether we use twirling operations in the intermediate steps or not, and we study them both analytically and numerically with Monte Carlo methods. When twirling is present, we are considering generalizations of the original distillation protocol. In the absence of twirling operations, we construct extensions of quantum privacy amplification algorithms valid for secure communications with qudits of any dimension D. When D is a prime number, we show that distillation protocols are optimal both qualitatively and quantitatively.

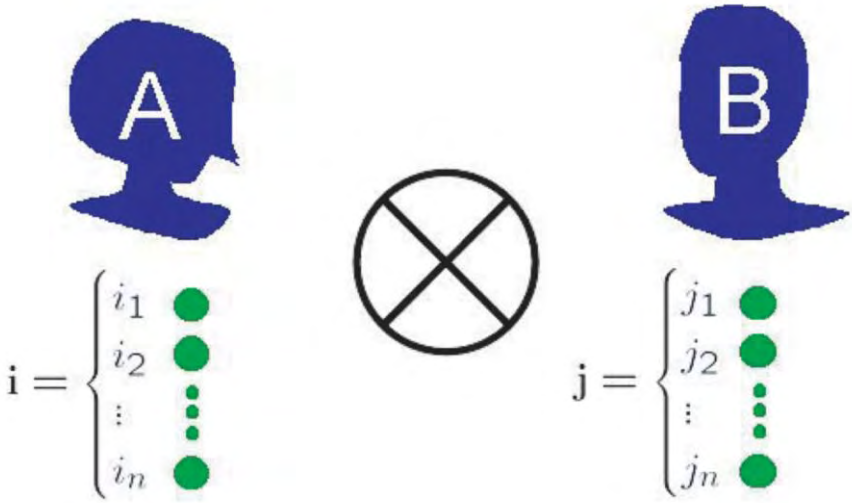


Figure 1. Alice and Bob share n Bell pairs of qudits. It is shown the pattern of indices in the tensor product of pairs.

2. Bell States for Qudits and the Module $\mathbf{Z}_D^n$

Why qudits after all? They are more resistant to noise than qubits. In quantum cryptography, they are more secure against eavesdropping attacks. Moreover, they violate more strongly local reality. A qudit is any quantum system of finite dimension D. The elements of a given orthogonal basis can be denoted $|x\rangle$ with $x = 0, \ldots, D-1$. This set of numbers is naturally identified with the elements of the set modulus

$$\mathbf{Z}_D := \mathbf{Z}/D\mathbf{Z}. \tag{1}$$

$\mathbf{Z}_D^n$ is a module and not a vector space. This is because $\mathbf{Z}_D$ is a field only for D prime. For example, for D = 4 we have:

$$2 \cdot (0, 2) = (0, 0). \tag{2}$$

The computational basis for n shared pairs and Bell (product) basis is [3]

$$|\mathbf{i}\ \mathbf{j}\rangle := \bigotimes_{k=1}^{n} |i_k\ j_k\rangle,$$

$$|\mathbf{i}\ \mathbf{j}\rangle_{\mathcal{B}} := \bigotimes_{k=1}^{n} |i_k\ j_k\rangle_{\mathcal{B}} = D^{-\frac{n}{2}} \sum_{\mathbf{k} \in \mathbf{Z}_D^n} \varphi(\mathbf{i} \cdot \mathbf{k}) |\mathbf{k}\ \mathbf{k} - \mathbf{j}\rangle, \tag{3}$$

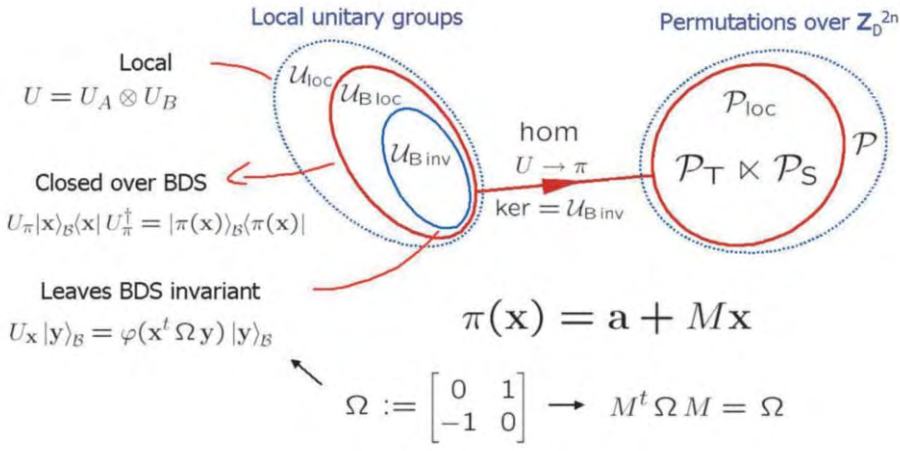


Figure 2. Schematic representation of the relationship among the unitary groups of local operations and the group of permutations in theorem 3.1.

where $\varphi(x) := \exp(\frac{2\pi i}{D} x)$ and $\mathbf{i}, \mathbf{j} \in \mathbf{Z}_D^n$. Bell diagonal states (BDS) are mixed states of the form

$$\rho^{(n)} = \sum_{\mathbf{x} \in \mathbf{Z}_D^{2n}} p_{\mathbf{x}}^{(n)} \, |\mathbf{x}\rangle_{\mathcal{B}} \langle \mathbf{x}|. \tag{4}$$

Fig. 1 pictures the notation of shared pairs by Alice and Bob.

In order to classify the anomalous vectors, consider the set of divisors of D:

$$\mathrm{div}(D) := \{\, d \in \mathbf{N} : d | D \,\}. \tag{5}$$

This set inherits the ordering of $\mathbf{N}$, and we shall use this property to introduce a suitable gcd function in $\mathbf{Z}_D$:

Definition 2.1 *For every $S \subset \mathbf{Z}_D$ we define the greatest common divisor of S, or* $\gcd(S)$*, to be the greatest $d \in \mathrm{div}(D)$ such that $\frac{D}{d} s = 0$, $\forall\, s \in S$.*

The nomenclature was chosen because for any $d \in \mathrm{div}(D)$ and $x \in \mathbf{Z}$ we have

$$d | x \iff D | \frac{D}{d} x \iff \frac{D}{d} x = 0 \pmod{D}, \tag{6}$$

and then for any set of integers X

$$\gcd(\overline{X}) = \max\{\, d \in \mathrm{div}(D) : d | x \,\forall\, x \in X \,\}, \tag{7}$$

where $\overline{X}$ is the corresponding set in $\mathbf{Z}_D$.

Vectors over $\mathbf{Z}_D$ are n-tuples of elements in $\mathbf{Z}_D$, and so we extend the gcd function to act over $\mathbf{Z}_D^n$ in the natural way, that is, if $\mathbf{v} = (v_1, \ldots v_n)$, $\gcd(\mathbf{v}) :=$

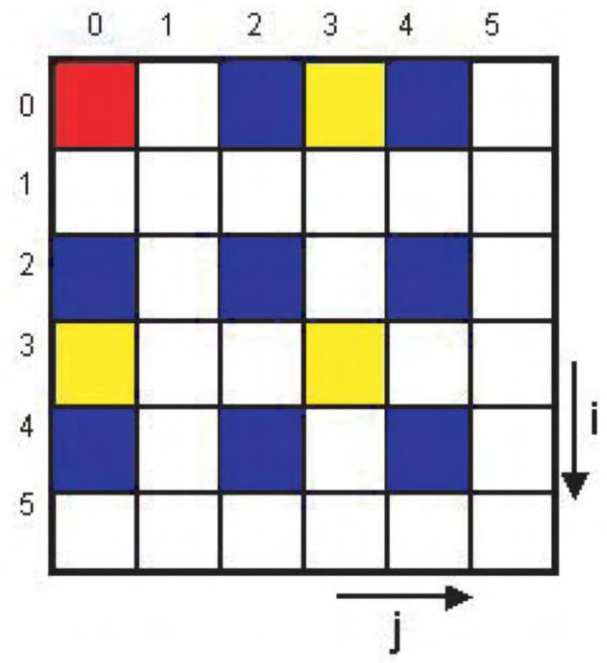


Figure 3. This is a representation of a generic heterotropic state (25) for $D = 6$ (single pair of qudits). Each cell represents the probability p_{ij} of a (Bell) diagonal element (4), with index given by the row and the column of the cell. The cells with the same colour have the same probability. There is one colour for each of the divisors of 6: 1 is associated to white, 2 to blue, 3 to yellow and 6 to red..

$\gcd(\{v_1, \ldots, v_n\})$. Now we can consider an equivalence relation in $\mathbf{Z}_D^n$ governed by the equality under the gcd function. The corresponding partition consists in the sets:

$$C_d(D, n) := \{\, \mathbf{v} \in \mathbf{Z}_D^n : \gcd(\mathbf{v}) = d \,\},\ d \in \mathrm{div}(D) \tag{8}$$

The most important of these sets is $C_1(D, n)$, since it contains those vectors $\mathbf{v}$ for with $\{\mathbf{v}\}$ is linearly independent. Later we will need its cardinality when considering properties of local unitary operators acting on diagonal Bell states. Thus, it is useful to define:

$$\phi_n(D) := \begin{cases} 1 & \text{if } D = 1 \\ \sharp C_1(D, n) & \text{if } D > 1 \end{cases} \tag{9}$$

For the particular case of $n = 1$, $\phi_1(x)$ corresponds to Euler's totient ϕ-function [4]. Euler's ϕ-function appears naturally in number theory since it gives for a natural number n, the cardinality of the set $\{\, m = 1, \ldots, n-1 : \gcd(m, n) = 1 \,\}$. That

is, $\phi_1(n)$ is the total number of coprime integers (or totatives) below or equal to n. For example, there are eight totatives of 24, namely, $\{1, 5, 7, 11, 13, 17, 19, 23\}$, thus $\phi_1(24) = 8$. For $n \neq 1$, we have therefore introduced a generalization of Euler's totient function for elements in $\mathbf{Z}_D^n$. The following lemma gives us how to compute the cardinalities of the sets $C_d(D, n)$, which shall naturally arise in our analysis of distillation protocols.

Lemma 2.2 *For every $n \in \mathbf{N}$, $D \in \mathbf{N} - \{1\}$ and $d \in \mathrm{div}(D)$:*

1.
$$\phi_n(D) = D^n \prod_{\substack{p|D \\ p \text{ prime}}} \frac{p^n - 1}{p^n} \tag{10}$$

2.
$$\sharp C_d(D, n) = \phi_n(\frac{D}{d}) \tag{11}$$

3.
$$\sum_{d' \in \mathrm{div}(D)} \phi_n(d') = D^n \tag{12}$$

The proof of this lemma can be found in [3].

3. Unitary Local Permutations on Bell Diagonal States

The main constraint Alice and Bob have to face when they intend to distill qudits is that they can perform only local operations. If we consider only unitary operations, we are led to the group $\mathcal{U}_{\mathrm{loc}}$ of local unitary operations. Its elements are all of the form

$$U = U_{\mathrm{A}} \otimes U_{\mathrm{B}}. \tag{13}$$

We study the subgroup $\mathcal{U}_{\mathrm{B\,loc}}$, defined as the group of local unitary operations which are closed over the space of Bell diagonal states (4).

Since the mapping of Bell states must be one to one, the action of any $U \in \mathcal{U}_{\mathrm{B\,loc}}(D, n)$ involves a permutation of the Bell states:

$$U \rho U^\dagger = \sum_{\mathbf{x} \in \mathbf{Z}_D^{2n}} p_{\mathbf{x}}^{(n)} \, |\pi(\mathbf{x})\rangle_{\mathcal{B}} \langle \pi(\mathbf{x})|, \tag{14}$$

where $\pi : \mathbf{Z}_D^{2n} \to \mathbf{Z}_D^{2n}$ is a permutation. So we introduce $\mathcal{P}_{\mathrm{loc}}(D, n)$, the *group of local permutations*, as the set of permutations over $\mathbf{Z}_D^{2n}$ implementable over Bell states by local (unitary) means.

We need to define several groups. Consider the family of unitary operators $u_{\mathbf{x}}$ ($\mathbf{x} \in \mathbf{Z}_D^{2n}$) over Bob's part of the system such that by definition

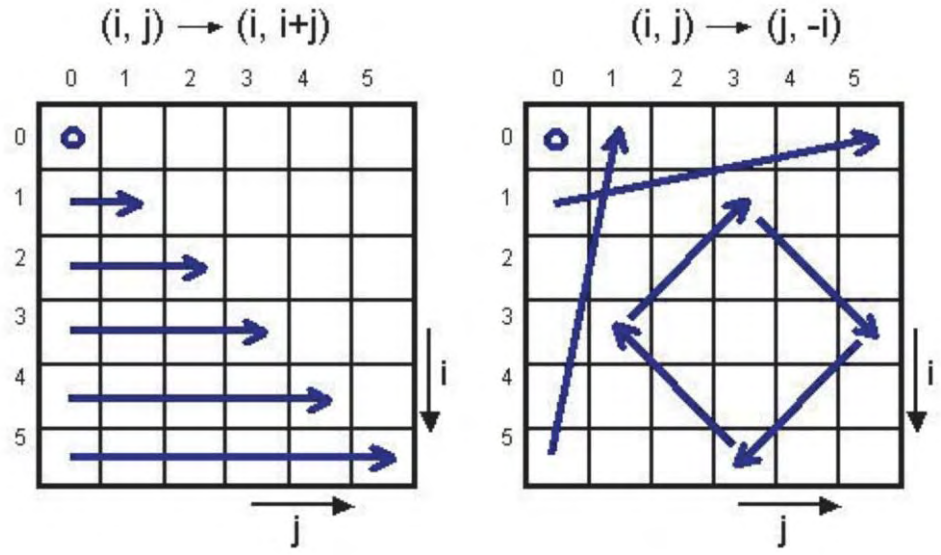


Figure 4. This is a representation of the two permutations generating the group $\mathcal{P}_{\mathrm{S}}$ for a pair of qudits. Each cell represents an element of $\mathbf{Z}_D^2$ and the arrows shows the action of the permutation.

$$1 \otimes u_{\mathbf{x}}^* \left|\mathbf{0}\right\rangle_{\mathcal{B}} := \left|\mathbf{x}\right\rangle_{\mathcal{B}}, \tag{15}$$

where conjugation is taken respect to the computational basis. With this operators at hand, we construct the group $\mathcal{U}_{\mathrm{B\,inv}}(D,n)$ with the elements $U_{\mathbf{x}} := u_{\mathbf{x}} \otimes {u_{\mathbf{x}}}^*$. We claim that it is a subgroup of $\mathcal{U}_{\mathrm{B\,loc}}(D,n)$. An explicit calculation shows that the action of its elements is

$$U_{\mathbf{x}} \left|\mathbf{y}\right\rangle_{\mathcal{B}} = \varphi(\mathbf{x}^t\, \Omega\, \mathbf{y}) \left|\mathbf{y}\right\rangle_{\mathcal{B}}, \tag{16}$$

where $\Omega \in \mathbf{M}_{2n\times 2n}(\mathbf{Z}_D)$ is

$$\Omega := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{17}$$

So the special feature of $\mathcal{U}_{\mathrm{B\,inv}}$ is that for $\rho^{(n)}$ Bell diagonal $U_{\mathbf{x}} \rho^{(n)} U_{\mathbf{x}}^\dagger = \rho^{(n)}$, which means that its elements implement the identity permutation.

We also define two subgroups of the group of permutations over $\mathbf{Z}_D^{2n}$. The *translation group* $\mathcal{P}_{\mathrm{T}}(D,n)$ contains the permutations of the form

$$\pi_{\mathbf{a}}(\mathbf{x}) = \mathbf{x} + \mathbf{a}, \tag{18}$$

with $\mathbf{a} \in \mathbf{Z}_D^{2n}$, and the *symplectic group* $\mathcal{P}_{\mathrm{S}}(D,n)$ contains in turn those whose action is

$$\pi_M(\mathbf{x}) = M\mathbf{x}, \tag{19}$$

where $M \in \mathbf{M}_{2n\times 2n}(\mathbf{Z}_D)$ is such that

$$M^t\,\Omega\,M = \Omega\,. \tag{20}$$

In Fig. 4 we picture the action of some of its generators. $\mathcal{P}_{\mathrm{S}}(D,n)$ is a finite non-simple group. A suitable generator set for this group is presented in [3].

Now, we can stablish our main result that plays an important role in the distillation protocols for qudits to be devised later on.

Theorem 3.1

1. *$\mathcal{P}_{\mathrm{loc}}$ is the semidirect product of $\mathcal{P}_{\mathrm{S}}$ and $\mathcal{P}_{\mathrm{T}}$:*

$$\mathcal{P}_{\mathrm{loc}}(D,n) = \mathcal{P}_{\mathrm{T}}(D,n) \ltimes \mathcal{P}_{\mathrm{S}}(D,n) \tag{21}$$

2. *Let h be the natural homomorphism from $\mathcal{U}_{\mathrm{B\,loc}}$ onto $\mathcal{P}_{\mathrm{loc}}$, then its kernel is*

$$\ker h = \mathcal{U}_{\mathrm{B\,inv}}(D,n) \otimes U(1), \tag{22}$$

where $U(1)$ denotes the global phase.

We prove this theorem in [3]. This theorem is summarized in Fig. 2.

4. Invariant States under Local Permutations: Heterotropic States

The twirling operation or depolarizing channel is defined for finite groups as follows:

$$T_{\mathcal{U}}(\rho) := \frac{1}{\sharp\mathcal{U}} \sum_{U\in\mathcal{U}} U\rho\,U^{\dagger}. \tag{23}$$

The group $\mathcal{P}_{\mathrm{S}}$ can also be successfully used in twirling operations. For ρ Bell diagonal states, we have:

$$T_{\mathcal{P}_{\mathrm{S}}}(\rho) := \frac{1}{\sharp\mathcal{P}_{\mathrm{S}}} \sum_{\pi\in\mathcal{P}_{\mathrm{S}}} U(\pi)\rho\,U(\pi)^{\dagger}. \tag{24}$$

To perform the sum we need to know which are the states invariant under the action of the group.

Theorem 4.1 *For every $\mathbf{x},\mathbf{y} \in \mathbf{Z}_D^{2n}$, $\gcd(\mathbf{x}) = \gcd(\mathbf{y})$ if and only if there exists a permutation in $\mathcal{P}_{\mathrm{S}}(D,n)$ with associated matrix M such that $\mathbf{x} = M\mathbf{y}$.*

We prove this theorem in [3].

Now, let us recall the partition in $\mathbf{Z}_D^{2n}$ associated to the function gcd (see (8)). We define the related states as heterotropic states (see Fig. 3)

$$\rho_d := \frac{1}{\sharp C_d} \sum_{\mathbf{x} \in C_d} |\mathbf{x}\rangle_{\mathcal{B}}\langle\mathbf{x}|. \tag{25}$$

These are the invariant states we were searching for. Thus, if ρ is Bell diagonal:

$$T_{\mathcal{P}_{\mathrm{S}}}(\rho) = \sum_{d \in \mathrm{div}(D)} \frac{\mathrm{Tr}(\rho_d \rho)}{\mathrm{Tr}(\rho_d \rho_d)} \rho_d \tag{26}$$

As a corollary, if D is prime there are just two Bell diagonal invariant states

$$\rho_1 = \frac{1}{D^{2n} - 1} \big(1 - |\mathbf{0}\rangle_{\mathcal{B}}\langle\mathbf{0}|\big), \tag{27}$$

$$\rho_D = |\mathbf{0}\rangle_{\mathcal{B}}\langle\mathbf{0}|, \tag{28}$$

and thus the result of the twirling operation is an isotropic state, which is the simplest example of an heterotropic state.

5. Permutation Based Distillation

In the distillation protocols we consider, which are iterative, each iteration cycle can be decomposed in the following steps:

1. *At start, Alice and Bob share n qudit pairs of dimension D and state matrix $\rho^{(n)}$.*
2. *They apply by local means one of the permutations $\pi_M \in \mathcal{P}_{\mathrm{S}}(D, n)$ in* (19).
3. *They measure the last $n-m$ qudit pairs, both of them in their computational basis.*
4. *If the results of the measurement agree for each of the measured pairs, they keep the first m pairs (in the state $\rho^{(m)}$). Else, they discard them.*

Fig. 7 shows a representation of the iteration cycle of a generic distillation protocol.

In most situations, the initial n pairs are independent and have equal state matrices ρ. In these cases

$$\rho^{(n)} = \rho^{\otimes n}. \tag{29}$$

In general (for $m > 1$) this does not guarantee that $\rho^{(m)}$ will be a product state, however, and thus it is preferable to consider the most general case.

Primes are special due to the appearance of one undesired attractor in the distillation process for each of the proper divisors of D. These attractors are related to invariant states under $\mathcal{P}_{\mathrm{loc}}(D, n)$.

Protocols with $m > 1$ are useless for low fidelities, but useful in other domains [3].

The recursion relations associated to these distillation protocols can be computed analitically and do not depend on non diagonal elements:

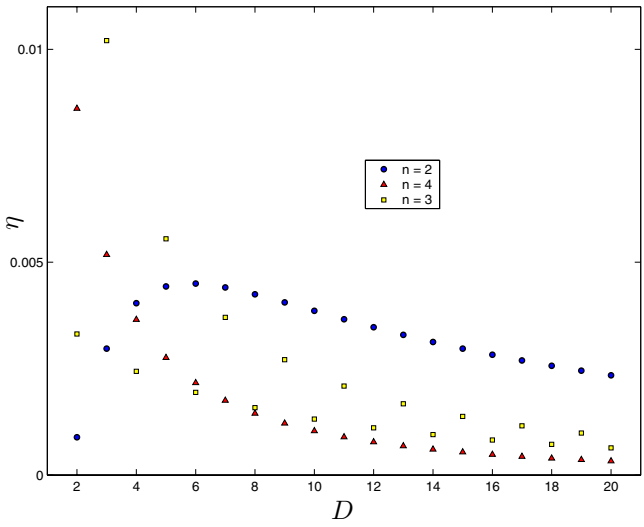


Figure 5. Values of the coefficient η for the considered twirled-assisted protocols with $n = 2, 3, 4$. Initial fidelity is close to $\frac{1}{D}$.

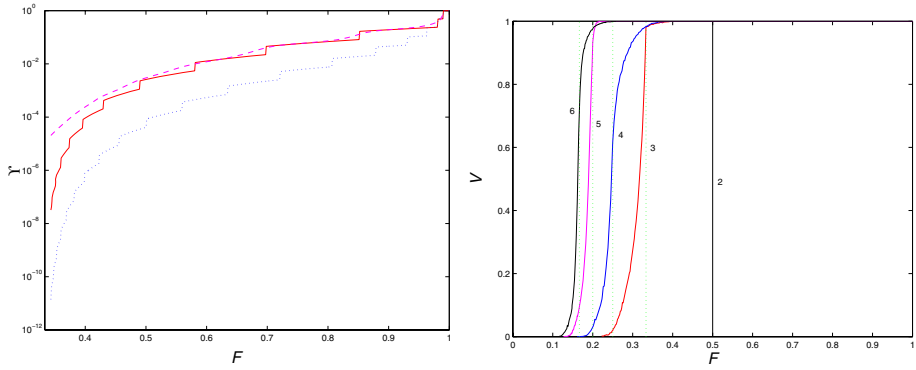


Figure 6. Left: Comparison of the yield respect to the twirling based protocol ($D = 3$). Right: Normalized volume of distilled Bell diagonal states. Non prime numbers behave much worse.

If we call V_M the space generated by the last $n - m$ rows of M (the matrix associated to π_M) the probability of obtaining the desired measure is

$$P = \sum_{\mathbf{x} \in V_M^{\perp}} p_{\mathbf{x}}^{(n)}, \tag{30}$$

and the recurrence relation for the probabilities is

$$p_{\mathbf{x}}^{(m)} = \frac{1}{P} \sum_{\mathbf{y} \in V_M} p_{\Omega\, \mathbf{y} + M^{-1} \bar{\mathbf{x}}}^{(n)}, \tag{31}$$

where $\mathbf{x} \in \mathbf{Z}_D^{2m}$ and $\bar{\mathbf{x}} \in \mathbf{Z}_D^{2n}$ is

$$\bar{\mathbf{x}} := (x_1, \ldots, x_m, \underbrace{0, \ldots, 0}_{n-m}, x_{n+1}, \ldots, x_{n+m}, \underbrace{0, \ldots, 0}_{n-m}). \tag{32}$$

We must distinguish between twirling assisted protocols and protocols without twirling.

For twirling assisted protocols, before each step, states are made isotropic using a suitable twirling operation. This reduces greatly the number of parameters describing the protocol. Low fidelity states cannot be distilled using hashing [5]. They deserve special attention. It is posible to define the joint performance η, which allows a direct comparison of protocols. This is shown in Fig. 5 [3].

For protocols without twirling, the Quantum Privacy Amplification algorithm [2] can be generalized to qudits by carefully choosing the permutation $\pi \in \mathcal{P}_{\text{loc}}(D, n)$ [3]. We get the following recursion relation:

$$p'_{ij} = \frac{1}{P} \sum_{k \in \mathbf{Z}_D} p_{i+k,-i-j-k}\, p_{k,j-k}, \tag{33}$$

$$P = \frac{1}{D} \sum_{\tilde{k} \in \mathbf{Z}_D} {p_{\tilde{k}\tilde{k}}}^2. \tag{34}$$

Numerical results are shwon in Fig. 6 [3].

References

[1] C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin, W.K. Wootters, "Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels", Phys. Rev. Lett. **76**, 722-725 (1996).

[2] D. Deutsch, A.Ekert, R. Jozsa, C. Macchiavello, S. Popescu, A. Sanpera, "Quantum Privacy Amplification and the Security of Quantum Cryptography over Noisy Channels", Phys. Rev. Lett. **77**, 2818-2821 (1996).

[3] H. Bombin, M.A. Martin-Delgado,
"Entanglement Distillation Protocols and Number Theory"; quant-ph/0503013. Phys. Rev. **A** accepted.

[4] J.H. Conway, R.K. Guy, "Euler's Totient Numbers." in *The Book of Numbers.* New York: Springer-Verlag, pp. 154-156, 1996.

[5] K. G. H. Vollbrecht and M. M. Wolf, "Efficient distillation beyond qubits", Phys. Rev. A **67**, 012303 (2003)

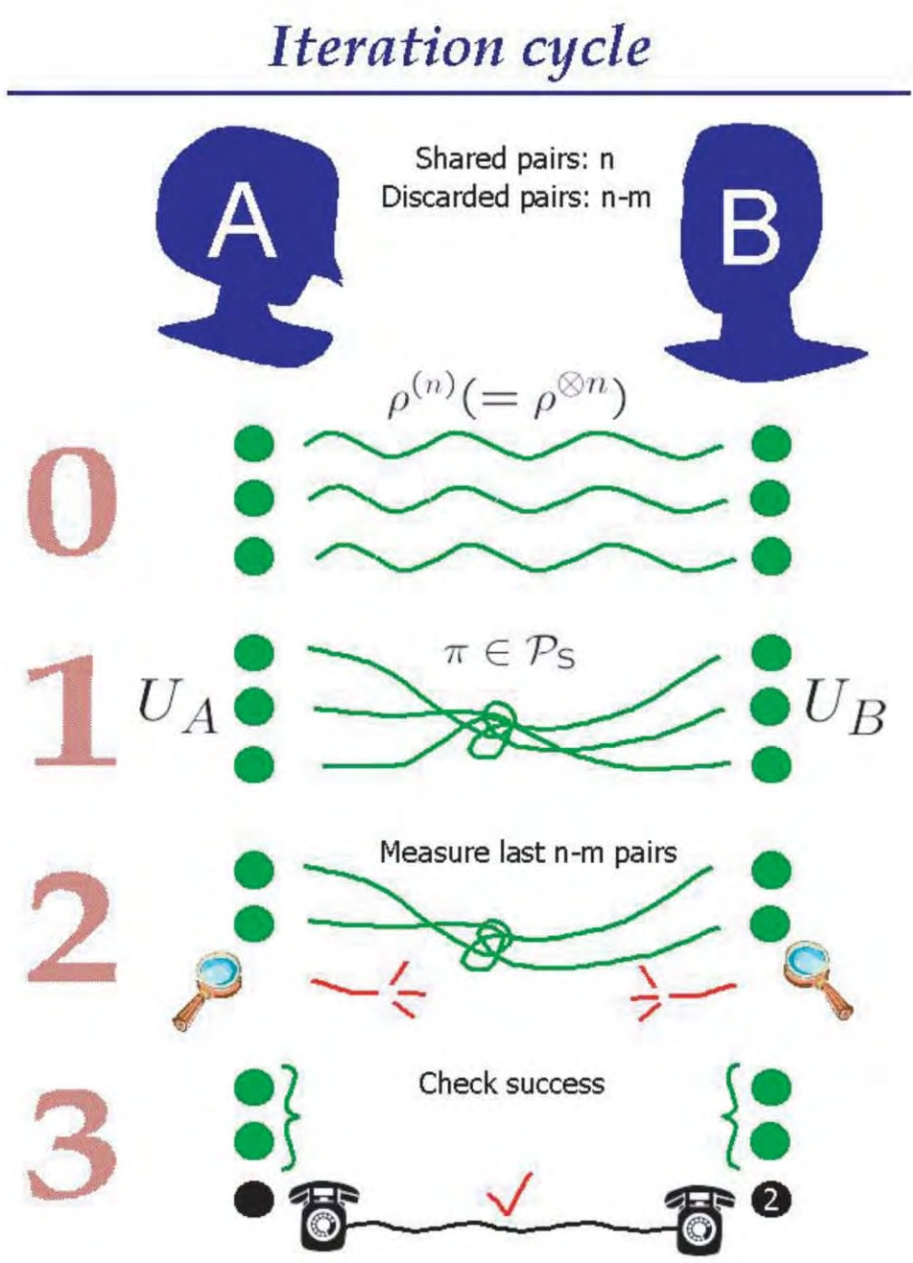


Figure 7. Pictorial represenation of the iteration cycle of a generic distillation protocol based on the local permutations of the group $\mathcal{P}_{\mathrm{loc}}(D,n)$.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Quantum Communication Channels in Infinite Dimensions

Holevo A. S. [1]

Abstract. These lectures are devoted to study of separability, entropy and channels in infinite dimensional Hilbert spaces. The first two sections deal with criteria of compactness for subsets in state space and collections of state ensembles. Next a general integral representation for separable states in the tensor product of infinite dimensional Hilbert spaces is given and an example of separable state that is not countably decomposable is provided. The structure theorem for the quantum communication channels that are entanglement-breaking is proved, generalizing the finite-dimensional result of M. Horodecki, Ruskai and Shor. We then discuss (dis)continuity properties of the quantum entropy and relative entropy in the infinite dimensional case. Sufficient conditions for the continuity of the output entropy of a quantum channel and related important entropic quantities are given. In case of compact constraints conditions for existence of optimal ensembles achieving the χ-capacity are given. These results are then applied to Gaussian channels.

Keywords. Hilbert space, compactness, separable state, entanglement, entropy, quantum communication channel

1. Introduction

These lectures are devoted to study of separability, entropy and channels in infinite dimensional Hilbert spaces. While major attention in quantum information theory up to now was paid to finite dimensional systems, there is an important and interesting class of Bosonic Gaussian systems which live in infinite dimensional Hilbert space. Although many questions for such systems with finite number of modes can be solved with finite dimensional matrix techniques, a general underlying Hilbert space operator analysis is indispensable. In general, any case involving spatial degrees of freedom requires consideration in infinite dimensional Hilbert space.

There are two important features essential for channels in infinite dimensions. One is the necessity of the input constraints (such as mean energy constraint for Gaussian channels) which prevent from infinite capacities. Another is the natural appearance of infinite, and, in general, continuous state ensembles understood as probability measures on the set of all quantum states. A general property which makes infinite dimensional objects close to finite dimensional is compactness. Therefore we devote the first two sections to criteria of compactness for subsets in state space and collections of ensembles.

Next a general integral representation for separable states in the tensor product of infinite dimensional Hilbert spaces is given and an example of separable states that are

[1]Correspondence to: Holevo A. S., Steklov Mathematical Institute, Gubkina 8, 119991 Moscow, Russia. Tel.: +7 095 1352291; Fax: +7 095 1350555; E-mail: holevo@mi.ras.ru.

not countably decomposable is provided. The structure theorem for the quantum communication channels that are entanglement-breaking is proved, generalizing the finite-dimensional result of M. Horodecki, Ruskai and Shor.

We then discuss (dis)continuity properties of the quantum entropy and relative entropy in the infinite dimensional case. Sufficient conditions for the continuity of the output entropy of a quantum channel and related important entropic quantities are given. In case of compact constraints we give conditions for existence of optimal ensembles achieving the χ -capacity. These results are then applied to the Gaussian channels.

Most of the new results presented in these lectures were obtained in close collaboration with M. E. Shirokov and R. F. Werner, see [9,10].

2. Convergence and Compactness in the State Space

In what follows $\mathcal{H}, \mathcal{K}, \ldots$ denote separable Hilbert spaces; $\mathfrak{T}(\mathcal{H})$ denotes the Banach space of trace-class operators in $\mathcal{H}$ with the trace norm $\|A\|_1 = \mathrm{Tr}|A|$, and $\mathfrak{S}(\mathcal{H})$ – the convex subset of all density operators. We shall also call them *states* for brevity, having in mind that a density operator ρ uniquely determines a normal state on the algebra of all bounded operators in $\mathcal{H}$. Equipped with the trace norm distance, $\mathfrak{S}(\mathcal{H})$ is a complete separable metric space which is not locally compact unless $\mathcal{H}$ is finite dimensional[2]. The set of extreme points $\mathrm{ext}\mathfrak{S}(\mathcal{H}) \equiv \mathfrak{P}(\mathcal{H})$ called *pure states* is a trace norm closed subset of $\mathfrak{S}(\mathcal{H})$ as follows from the condition $\rho^2 = \rho$ characterizing the pure states.

Theorem 1. [3], [4] *Let ρ_n be a sequence of density operators converging to a density operator ρ in the weak operator topology. Then ρ_n converges to ρ in the trace norm.*

Proof. Let $\{\rho_n\}$ be a sequence of density operators in $\mathcal{H}$ such that $\langle\psi|\,\rho_n|\phi\rangle \to \langle\psi|\,\rho|\phi\rangle$ as $n \to \infty$ for all $\phi, \psi \in \mathcal{H}$, then we have to show that $\|\rho_n - \rho\|_1 = \mathrm{Tr}|\rho_n - \rho| \to 0$. Indeed, for any finite-dimensional projector P,

$$\|\rho_n - \rho\|_1 \le \|P(\rho_n - \rho)P\|_1 + 2\,\|P\rho_n(I-P)\|_1 + 2\,\|P\rho(I-P)\|_1 + \|(I-P)\rho_n(I-P)\|_1 + \|(I-P)\rho(I-P)\|_1 .$$

The first term on the right tends to zero for any choice of P, since $P\rho_n P \to P\rho P$ due to the weak convergence, and to equivalence of all types of convergence in the finite-dimensional case. For the last two terms we have:

$$\|(I-P)\rho(I-P)\|_1 = \mathrm{Tr}(I-P)\rho(I-P) = \mathrm{Tr}(I-P)\rho = 1 - \mathrm{Tr}P\rho,$$

which can be made arbitrarily small by the choice of P, and

$$\|(I-P)\rho_n(I-P)\|_1 = 1 - \mathrm{Tr}P\rho_n \to 1 - \mathrm{Tr}P\rho$$

by the weak convergence. Finally, the intermediate terms can be evaluated as follows:

[2] In the mathematical literature one usually includes infinite dimensionality in the definition of the Hilbert space; its finite dimensional prototype is called the *unitary space*.

$$\|P\rho_n(I-P)\|_1 = \mathrm{Tr}U^*P\rho_n(I-P) = \mathrm{Tr}U^*P\sqrt{\rho_n}\sqrt{\rho_n}(I-P),$$

where U is the unitary operator from the polar decomposition of $P\rho_n(I-P)$. Then by the operator Cauchy-Schwarz inequality for the trace

$$\begin{aligned}\mathrm{Tr}U^*P\sqrt{\rho_n}\sqrt{\rho_n}(I-P) &\le \sqrt{\mathrm{Tr}U^*P\sqrt{\rho_n}\sqrt{\rho_n}PU}\sqrt{\mathrm{Tr}(I-P)\sqrt{\rho_n}\sqrt{\rho_n}(I-P)} \\ &= \sqrt{\mathrm{Tr}P\rho_n}\sqrt{1-\mathrm{Tr}P\rho_n} \to \sqrt{\mathrm{Tr}P\rho}\sqrt{1-\mathrm{Tr}P\rho}\end{aligned}$$

which again can be made small by the choice of P.□

One can consider this theorem as a noncommutative generalization of the fact that weak convergence of probability distributions on a discrete probability space implies l_1 convergence. By using this theorem we can amplify to the trace norm topology the weak compactness criterion given in [16] under the name "the noncommutative Prokhorov's theorem". Recall that compactness of a subset $\mathcal{A}$ in a complete separable metric space means that any sequence $\{\rho_n\} \subset \mathcal{A}$ contains a subsequence converging to $\rho \in \mathcal{A}$.

Theorem 2. *A trace norm closed subset $\mathcal{A}$ of $\mathfrak{S}(\mathcal{H})$ is compact if and only if for arbitrary $\varepsilon > 0$ there exists a finite rank projector P_ε such that $\mathrm{Tr}P_\varepsilon\rho > 1-\varepsilon$ for all $\rho \in \mathcal{A}$.*

Proof. Let $\mathcal{A}$ be a trace norm compact subset of $\mathfrak{S}(\mathcal{H})$. Suppose that there exists $\varepsilon > 0$ such that for arbitrary finite rank projector P there is a state $\rho \in \mathcal{A}$ such that $\mathrm{Tr}P\rho \le 1-\varepsilon$. Let P_n be a sequence of finite rank projectors in $\mathcal{H}$ monotonously converging to the identity operator $I_{\mathcal{H}}$, and ρ_n be the corresponding sequence of states in $\mathcal{A}$ such that $\mathrm{Tr}P_n\rho_n \le 1-\varepsilon$. By compactness of $\mathcal{A}$ there exists a subsequence ρ_{n_k} converging to a state $\rho_* \in \mathcal{A}$. By construction $\mathrm{Tr}P_{n_l}\rho_{n_k} \le \mathrm{Tr}P_{n_k}\rho_{n_k} \le 1-\varepsilon$ for $k > l$. Hence

$$\mathrm{Tr}\rho_* = \lim_{l\to+\infty}\mathrm{Tr}P_{n_l}\rho_* = \lim_{l\to+\infty}\lim_{k\to+\infty}\mathrm{Tr}P_{n_l}\rho_{n_k} \le 1-\varepsilon,$$

which contradicts to the fact that $\rho_* \in \mathcal{A} \subseteq \mathfrak{S}(\mathcal{H})$.

Conversely, let $\mathcal{A}$ be a subset of $\mathfrak{S}(\mathcal{H})$ satisfying the criterion. Let ρ_n be an arbitrary sequence in $\mathcal{A}$. Since the unit ball in $\mathfrak{B}(\mathcal{H})$ is compact in the weak operator topology [2], there exists a subsequence ρ_{n_k} converging to a positive operator ρ_* in this topology. We have

$$\mathrm{Tr}\rho_* \le \liminf_{k\to\infty}\mathrm{Tr}\rho_{n_k} = 1,$$

therefore to prove that ρ_* is a state it is sufficient to show that $\mathrm{Tr}\rho_* \ge 1$. Let $\varepsilon > 0$ and P_ε be the corresponding projector. We have

$$\mathrm{Tr}\rho_* \ge \mathrm{Tr}P_\varepsilon\rho_* = \lim_{k\to\infty}\mathrm{Tr}P_\varepsilon\rho_{n_k} > 1-\varepsilon,$$

where the equality follows from the fact that P_ε has finite rank. Thus $\mathrm{Tr}\rho_* = 1$ and ρ_* is a state. Theorem 1 implies that the subsequence ρ_{n_k} converges to the state ρ_* in the trace norm. Thus the set $\mathcal{A}$ is trace norm compact. □

An unbounded positive operator E in $\mathcal{H}$ with discrete spectrum of finite multiplicity will be called *$\mathfrak{H}$-operator*. Let Q_n be the spectral projector of E corresponding to the lowest n eigenvalues. Following [8] we shall denote

$$\mathrm{Tr}\rho E = \lim_{n\to\infty} \mathrm{Tr}\rho Q_n E, \tag{1}$$

where the sequence on the right side is monotonously nondecreasing, hence has a limit (possibly infinite).

Theorem 3. *A closed subset $\mathcal{A} \subseteq \mathfrak{S}(\mathcal{H})$ is compact if and only if there exist an $\mathfrak{H}$-operator E and a positive number h such that $\mathrm{Tr}\rho E \leq h$ for all $\rho \in \mathcal{A}$.*

Proof. Let E be an $\mathfrak{H}$-operator, λ_n its largest eigenvalue in the subspace $Q_n(\mathcal{H})$, so that $\lambda_n \uparrow +\infty$. Since

$$\mathrm{Tr}\rho Q_n > 1 - \lambda_{n+1}^{-1}\mathrm{Tr}\rho E \geq 1 - \lambda_{n+1}^{-1}h,$$

the set $\{\rho : \mathrm{Tr}\rho E \leq h\}$ is a compact subset of $\mathfrak{S}(\mathcal{H})$ by theorem 2.

Conversely, by theorem 2, for any $n = 1, 2, \ldots$ there exists a finite rank projector P_n such that $\mathrm{Tr}\rho P_n \geq 1 - n^{-3}$ for all ρ in $\mathcal{A}$. Without loss of generality we may assume that $\bigvee_{k=1}^{+\infty} P_k(\mathcal{H}) = \mathcal{H}$, where $\bigvee$ denotes closed linear span of the subspaces. Let $\hat{P}_n$ be the projector on the finite dimensional subspace $\bigvee_{k=1}^{n} P_k(\mathcal{H})$. Then $\hat{P}_{n+1} \geq \hat{P}_n)$ and $E = \sum_{n=1}^{+\infty} n(\hat{P}_{n+1} - \hat{P}_n)$ is an $\mathfrak{H}$-operator satisfying

$$\mathrm{Tr}\rho E = \sum_{n=1}^{+\infty} n\mathrm{Tr}\rho(\hat{P}_{n+1} - \hat{P}_n) \leq \sum_{n=1}^{+\infty} n\mathrm{Tr}\rho(I_{\mathcal{H}} - \hat{P}_n) \leq \sum_{n=1}^{+\infty} n^{-2} = h$$

for arbitrary state ρ in the set $\mathcal{A}$. □

3. Convergence and Compactness for State Ensembles

An ensemble of states will be represented by a Borel probability measure on $\mathfrak{S}(\mathcal{H})$. This coincides with the usual notion of ensemble as a finite collection of states ρ_i taken with probabilities π_i in case the measure has finite support. The relation

$$\bar{\rho}(\pi) = \int_{\mathfrak{S}(\mathcal{H})} \sigma\pi(d\sigma), \tag{2}$$

where the integral exists as the Bochner integral by theorem 1, defines a state called *barycenter* of measure π (or the *average* of the ensemble).

We shall denote by $\mathcal{P}_{\mathfrak{S}(\mathcal{H})}$ the convex set of all Borel probability measures on $\mathfrak{S}(\mathcal{H})$. Recall that weak convergence in $\mathcal{P}_{\mathfrak{S}(\mathcal{H})}$ means convergence of integrals of all continuous bounded functions (see e. g. [15]).

Lemma 1. *The map $\pi \mapsto \bar{\rho}(\pi)$ is continuous.*

Proof. Let $\pi_n \to \pi$ weakly. For arbitrary $\varphi, \psi \in \mathcal{H}$ the matrix element $\rho \mapsto \langle\varphi|\rho\psi\rangle$ is bounded continuous function hence

$$\int \rho \mapsto \langle\varphi|\rho\psi\rangle\pi_n(d\rho) \to \int \rho \mapsto \langle\varphi|\rho\psi\rangle\pi(d\rho),$$

which implies that $\bar{\rho}(\pi_n) \to \bar{\rho}(\pi)$ in the weak operator topology, and by theorem 1, in the trace norm. □

Theorem 4. *A subset $\mathcal{P} \subseteq \mathcal{P}_{\mathfrak{S}(\mathcal{H})}$ is weakly compact if and only if its image $\mathcal{A}$ under the map $\pi \mapsto \bar{\rho}(\pi)$ given by (2) is a compact subset of $\mathfrak{S}(\mathcal{H})$.*

Proof. The set $\mathcal{A}$ is the image of the set $\mathcal{P}$ under the continuous map $\pi \mapsto \bar{\rho}(\pi)$, hence it is compact.

Conversely, let the set $\mathcal{A}$ be compact. By theorem 3 there exists an $\mathfrak{H}$-operator E such that $\mathrm{Tr}\rho E \leq h$ for all ρ in $\mathcal{A}$. For arbitrary $\pi \in \mathcal{P}$ we have

$$\int\limits_{\mathfrak{S}(\mathcal{H})} (\mathrm{Tr}\rho E)\pi(d\rho) = \mathrm{Tr}\left(\int\limits_{\mathfrak{S}(\mathcal{H})} \rho\pi(d\rho)\ E\right) = \mathrm{Tr}\bar{\rho}(\pi)E \leq h \tag{3}$$

The existence of the integral on the left hand side and the first equality follow from monotonous convergence theorem, since by (1) the function $\mathrm{Tr}\rho E$ is the limit of nondecreasing sequence of continuous bounded functions $\mathrm{Tr}\rho Q_n E$.

The set $\mathcal{P}$ is weakly closed as inverse image of closed set $\mathcal{A}$ under the continuous map $\pi \mapsto \bar{\rho}(\pi)$. Let $\mathcal{A}_\varepsilon = \{\rho : \mathrm{Tr}\rho E \leq h\varepsilon^{-1}\}$. The set $\mathcal{A}_\varepsilon$ is a compact subset of $\mathfrak{S}(\mathcal{H})$ for any ε by theorem 3. By (3) for any measure π in $\mathcal{P}$ we have

$$\pi(\mathfrak{S}(\mathcal{H})\backslash\mathcal{A}_\varepsilon) = \int\limits_{\mathfrak{S}(\mathcal{H})\backslash\mathcal{A}_\varepsilon} \pi(d\rho) \leq \varepsilon h^{-1} \int\limits_{\mathfrak{S}(\mathcal{H})\backslash\mathcal{A}_\varepsilon} (\mathrm{Tr}\rho E)\pi(d\rho) \leq \varepsilon \tag{4}$$

By Prokhorov's theorem [15] the closed set $\mathcal{P}$ is compact.□

The following lemma is an amplified version of the Choquet decomposition [1] adapted to the case of closed convex subsets of $\mathfrak{S}(\mathcal{H})$. We denote by $\mathrm{co}\mathcal{A}(\overline{\mathrm{co}}\mathcal{A})$ the convex hull (closure) of a set $\mathcal{A}$ defined as collection of all convex combinations (and their trace norm limits) of elements of $\mathcal{A}$.

Theorem 5. *Let $\mathcal{A}$ be a closed subset of $\mathfrak{S}(\mathcal{H})$. Then $\overline{\mathrm{co}}\mathcal{A}$ coincides with the set of barycenters of all Borel probability measures supported by $\mathcal{A}$.*

Proof. Let $\rho_0 \in \overline{\mathrm{co}}\mathcal{A}$. Then there is a sequence $\{\rho_n\} \subseteq \mathrm{co}\mathcal{A}$ converging to ρ_0, so that $\{\rho_n\}$ is relatively compact in $\mathfrak{S}(\mathcal{H})$. The density operator ρ_n is barycenter of Borel probability measure π_n finitely supported on $\mathcal{A}$. By the compactness criterion of theorem 4, the sequence $\{\pi_n\}$ is weakly relatively compact and thus has a partial limit π_0, which is also supported by the set $\mathcal{A}$ due to theorem 6.1 in [15]. Lemma 1 implies that the state ρ_0 is the barycenter of the measure π_0.

Conversely, let π be an arbitrary probability measure supported by $\mathcal{A}$. By theorem 6.3 in [15] this measure can be weakly approximated by a sequence of measures π_n finitely supported by $\mathcal{A}$. Since $\bar{\rho}(\pi_n)$ is in $\mathrm{co}\mathcal{A}$ for all n, we conclude that $\bar{\rho}(\pi)$ is in $\overline{\mathrm{co}}\mathcal{A}$ due to continuity of the map $\pi \mapsto \bar{\rho}(\pi)$. □

4. Separable States

Definition 1. *A state in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ is called separable if it is in the convex closure of the set of all product states in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$.*

Lemma 1 implies that separable states are precisely those states which admit the representation

$$\rho = \int_{\mathfrak{S}(\mathcal{H})} \int_{\mathfrak{S}(\mathcal{K})} (\rho_{\mathcal{H}} \otimes \rho_{\mathcal{K}})\, \mu(d\rho_{\mathcal{H}} d\rho_{\mathcal{K}}), \tag{5}$$

where μ is a Borel probability measure on $\mathfrak{S}(\mathcal{H}) \times \mathfrak{S}(\mathcal{K})$. In the finite dimensional case application of Caratheodory's theorem reduces this to the familiar definition of separable state as finite convex combination of pure product states. If for a separable state ρ it is possible to find a representation (5) with purely atomic μ, we call the state *countably decomposable*. A necessary condition for this is existence of nonzero vectors $|\alpha\rangle \in \mathcal{H}, |\beta\rangle \in \mathcal{K}$ such that

$$\rho \geq |\alpha\rangle\langle\alpha| \otimes |\beta\rangle\langle\beta|, \tag{6}$$

cf. [20]. We shall show that there are separable states which do not satisfy this condition and hence are not countably decomposable.

In the definition 1 one can replace the set of all product states by the set of all products of pure states. It is known that the subset $\mathfrak{P}(\mathcal{H})$ of pure states (extreme points of $\mathfrak{S}(\mathcal{H})$) is closed in the trace-norm topology. Theorem 5 then implies that a state ρ is separable if and only if there is a Borel measure ν on $\mathfrak{P}(\mathcal{H}) \times \mathfrak{P}(\mathcal{K})$ such that, with some abuse of notation,

$$\rho = \int_{\mathfrak{P}(\mathcal{H})} \int_{\mathfrak{P}(\mathcal{K})} |\varphi\rangle\langle\varphi| \otimes |\psi\rangle\langle\psi| \nu(d\varphi d\psi). \tag{7}$$

We shall consider the one-dimensional shift group $\mathbb{R}$. Let $\mathcal{H} = L^2(\mathbb{R})$ with the normalized Lebesgue measure $\frac{dx}{2\pi}$, and let

$$\langle p|\psi\rangle = \int e^{-ixp} \psi(x) \frac{dx}{2\pi} \tag{8}$$

be the momentum representation. Consider the unitary representation of the shift group $x \to V_x$, where $(V_u\psi)(x) = \psi(x-u)$.

Proposition 1. [10] *Let $p_0(x)$ be a bounded probability density on $\mathbb{R}$. For arbitrary state vectors $|\varphi_j\rangle \in \mathcal{H}_j \simeq L^2(\mathbb{R}); j = 1, 2$, with almost nowhere vanishing momentum representations (8) the separable state*

$$\rho_{12} = \int V_x^{(1)} |\varphi_1\rangle\langle\varphi_1| V_x^{(1)*} \otimes V_x^{(2)} |\varphi_2\rangle\langle\varphi_2| V_x^{(2)*}\, p_0(x) dx \tag{9}$$

in $\mathcal{H}_1 \otimes \mathcal{H}_2$ is not countably decomposable.

Proof. Suppose ρ is countably decomposable, then by (6) there exist nonzero $\alpha_j \in \mathcal{H}_j$ such that

$$\rho_{12} \geq |\alpha_1\rangle\langle\alpha_1| \otimes |\alpha_2\rangle\langle\alpha_2|. \tag{10}$$

Taking partial traces we obtain

$$\int V_x^{(j)} |\varphi_j\rangle\langle\varphi_j| V_x^{(j)*}\, p_0(x) dx \geq |\alpha_j\rangle\langle\alpha_j|,$$

whence

$$|\langle p|\varphi_j\rangle| \geq |\langle p|\alpha_j\rangle|; \quad j = 1, 2, \tag{11}$$

for almost all $p \in \mathbb{R}$.

Inequality (10) implies that

$$\int \left|\langle\lambda_1|V_x^{(1)}|\varphi_1\rangle\right|^2 \left|\langle\lambda_2|V_x^{(2)}|\varphi_2\rangle\right|^2 \; p_0(x)dx \geq |\langle\lambda_1|\alpha_1\rangle|^2|\langle\lambda_2|\alpha_2\rangle|^2 \tag{12}$$

for arbitrary $\lambda_j \in L^2(\mathbb{R})$. For technical convenience we will assume that the square integrable functions $p \mapsto \langle\lambda_j|p\rangle$ have finite support. Introducing

$$\mu_j(x) = \langle\lambda_j|V_x^{(j)}|\varphi_j\rangle = \int \langle\lambda_j|p\rangle\langle p|\varphi_j\rangle e^{ipx}dp,$$

so that $\langle\lambda_j|p\rangle = \frac{\langle p|\mu_j\rangle}{\langle p|\varphi_j\rangle}$, we see that

$$\mu_j \mapsto \langle\lambda_j|\alpha_j\rangle = \int \langle p|\mu_j\rangle \frac{\langle p|\alpha_j\rangle}{\langle p|\varphi_j\rangle}dp$$

are linear functionals of μ_j running over the dense subspace $\mathcal{L}$ consisting of functions in $L^2(\mathbb{R})$ with finite support in the momentum representation. These functionals are in fact continuous. Indeed, choosing λ_2 such that $\langle\lambda_2|\alpha_2\rangle \neq 0$, and using the Cauchy-Schwarz inequality for $|\mu_2(x)|^2 = \left|\langle\lambda_2|V_x^{(2)}|\varphi_2\rangle\right|^2$, we find from (12)

$$|\langle\lambda_1|\alpha_1\rangle|^2 \leq \sup_x p_0(x) \frac{\|\lambda_2\|^2 \|\varphi_2\|^2}{\langle\lambda_2|\alpha_2\rangle} \int |\mu_1(x)|^2 \; dx$$

for all $\mu_1 \in \mathcal{L}$. Hence by Riesz theorem there exists $\beta_1 \in L^2(\mathbb{R})$ such that $\langle\lambda_1|\alpha_1\rangle = \langle\beta_1|\mu_1\rangle$. Applying similar reasoning to $j = 2$, we can transform (12) to the form

$$|\langle\beta_1|\mu_1\rangle|^2|\langle\beta_2|\mu_2\rangle|^2 \leq \int |\mu_1(x)\mu_2(x)|^2 \; dx, \tag{13}$$

where $\beta_j \in L^2(\mathbb{R})$.

Now we can extend the inequality (13) to more general functions μ_j for which both sides of this inequality are defined, e. g. for almost everywhere uniformly bounded functions from $L^2(\mathbb{R})$. The characteristic functions of intervals belong to this class, and so is a dense set of functions with support in any specified interval.

Consider a partitioning of $\mathbb{R}$ into intervals of length $\leq \varepsilon$, and pick one of these intervals, say I_2, on which β_2 is not a.e. zero. Then we can find an admissible function μ_2 supported in I_2 such that $\langle\beta_2|\mu_2\rangle \neq 0$. But then, for any μ_1 supported on the complement of I_2, the right hand side of (13) vanishes, and therefore β_1 vanishes a.e. on the complement of I_2. It follows that the support of β_1 has measure $\leq \varepsilon$ for all ε, i.e., β_1 vanishes a.e., and hence $\alpha_1 = 0$. □

5. Entanglement-Breaking Channels

A *channel* is a linear map Φ:$\mathfrak{T}(\mathcal{H}) \mapsto \mathfrak{T}(\mathcal{H}')$ with the properties:

1) $\Phi(\mathfrak{S}(\mathcal{H})) \subseteq \mathfrak{S}(\mathcal{H}')$; this implies that Φ is bounded map and hence is uniquely determined by the infinite matrix $[\Phi(|i\rangle\langle j|)]$, where $\{|i\rangle\}$ is an orthonormal basis in $\mathcal{H}$.

2) The matrix $[\Phi(|i\rangle\langle j|)]$ is positive definite in the sense that for a collection of vectors $\{|\psi_i\rangle\} \subseteq \mathcal{H}$ with finite number of nonzero elements

$$\sum_{ij} \langle\psi_i|\Phi(|i\rangle\langle j|)|\psi_j\rangle \geq 0. \tag{14}$$

The condition 2) is equivalent to the fact that the dual channel Φ^* defined by the relation

$$\mathrm{Tr}\rho\Phi^*(X) = \mathrm{Tr}\Phi(\rho)X, \quad \rho \in \mathfrak{T}(\mathcal{H}, X \in \mathfrak{B}(\mathcal{H}),$$

where $\mathfrak{B}(\mathcal{H})$ is the algebra of all bounded operators in $\mathcal{H}$, is completely positive.

Definition 2. *A channel Φ is called entanglement-breaking if for arbitrary Hilbert space $\mathcal{K}$ and arbitrary state $\omega \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ the state $(\Phi \otimes \mathrm{Id}_{\mathcal{K}})(\omega)$, where $\mathrm{Id}_{\mathcal{K}}$ is the identity channel in $\mathfrak{S}(\mathcal{K})$, is separable in the sense of definition 1.*

Note that in this definition one can restrict to finite dimensional Hilbert spaces $\mathcal{K}$. Indeed, an arbitrary state ω in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ with an infinite dimensional $\mathcal{K}$ can be approximated by the states

$$\omega_n = (\mathrm{Tr}(I_{\mathcal{H}} \otimes Q_n)\omega)^{-1}(I_{\mathcal{H}} \otimes Q_n)\omega(I_{\mathcal{H}} \otimes Q_n),$$

where $\{Q_n\}$ is the sequence of the spectral projectors of the partial state $\mathrm{Tr}_{\mathcal{H}}\omega$ corresponding to its n largest eigenvalues. Each state ω_n can be considered as a state in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K}_n)$, where $\mathcal{K}_n = Q_n(\mathcal{K})$ is n-dimensional Hilbert space. If $(\Phi \otimes \mathrm{Id})(\omega_n)$ is separable for all n, then $(\Phi \otimes \mathrm{Id})(\omega)$ is also separable as a limit of sequence of separable states.

The following theorem is a generalization of the result in [12] to the infinite dimensional case.

Theorem 5. [10] *Channel Φ is entanglement-breaking if and only if there is a complete separable metric space $\mathcal{X}$, a Borel $\mathfrak{S}(\mathcal{H}')$-valued function $x \mapsto \rho'(x)$ and a positive operator-valued Borel measure (POVM) $M(dx)$ on $\mathcal{X}$ such that*

$$\Phi(\rho) = \int_{\mathcal{X}} \rho'(x)\mu_\rho(dx), \tag{15}$$

where $\mu_\rho(B) = \mathrm{Tr}\rho M(B)$ for all Borel $B \subseteq \mathcal{X}$.

Proof. Notice first that conditions 1),2) in the definition of channel are readily verified for the map (15). Let us show that the channel (15) is entanglement-breaking. Let $\omega \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$, where $\mathcal{K}$ is a finite dimensional Hilbert space. We have

$$(\Phi \otimes \mathrm{Id}_{\mathcal{K}})(\omega) = \int_{\mathcal{X}} \rho'(x) \otimes m_\omega(dx), \tag{16}$$

where

$$m_\omega(B) = \mathrm{Tr}_{\mathcal{H}}\omega(M(B)\otimes I_{\mathcal{K}}), \quad B \subseteq \mathcal{X}.$$

It is easy to see that any matrix element of m_ω (in a particular basis) is a complex valued measure on $\mathcal{X}$ absolutely continuous with respect to the probability measure $\mu_\omega(B) = \mathrm{Tr}m_\omega(B),\ B \subseteq \mathcal{X}$. The Radon-Nikodym theorem implies representation

$$m_\omega(B) = \int\limits_B \sigma_\omega(x)\mu_\omega(dx),$$

where $\sigma_\omega(x)$ is a function on $\mathcal{X}$ taking values in $\mathfrak{S}(\mathcal{K})$. By using this representation we can rewrite (16) as

$$(\Phi\otimes \mathrm{Id}_{\mathcal{K}})(\omega) = \int\limits_{\mathcal{X}} \rho'(x)\otimes\sigma_\omega(x)\mu_\omega(dx), \tag{17}$$

which reduces to (5) by change of variables and hence is separable by Lemma 1.

Conversely, let Φ be an entanglement-breaking channel. Fix a state σ in $\mathfrak{S}(\mathcal{H})$ of full rank and let $\{|i\rangle\}_{i=1}^{+\infty}$ be the basis of eigenvectors of σ with the corresponding (positive) eigenvalues $\{\lambda_i\}_{i=1}^{+\infty}$. Consider the vector

$$|\Omega\rangle = \sum_{i=1}^{+\infty} \lambda_i^{1/2}|i\rangle\otimes|i\rangle$$

in the space $\mathcal{H}\otimes\mathcal{H}$. Since Φ is entanglement-breaking, the state

$$\rho = (\mathrm{Id}_{\mathcal{H}}\otimes\Phi)(|\Omega\rangle\langle\Omega|) \tag{18}$$

in $\mathfrak{S}(\mathcal{H}\otimes\mathcal{H}')$ is separable. By (7) there exists a probability measure ν on $\mathfrak{P}(\mathcal{H})\times\mathfrak{P}(\mathcal{H}')$ such that

$$(\mathrm{Id}_{\mathcal{H}}\otimes\Phi)(|\Omega\rangle\langle\Omega|) = \int\limits_{\mathfrak{P}(\mathcal{H})}\int\limits_{\mathfrak{P}(\mathcal{H}')} |\varphi\rangle\langle\varphi|\otimes|\psi\rangle\langle\psi|\nu(d\varphi d\psi). \tag{19}$$

This implies

$$\begin{aligned}\sigma &= \mathrm{Tr}_{\mathcal{H}'}(Id_{\mathcal{H}}\otimes\Phi)(|\Omega\rangle\langle\Omega|)\\ &= \int\limits_{\mathfrak{P}(\mathcal{H})}\int\limits_{\mathfrak{P}(\mathcal{H}')} |\varphi\rangle\langle\varphi|\nu(d\varphi d\psi)\\ &= \int\limits_{\mathfrak{P}(\mathcal{H})}\int\limits_{\mathfrak{P}(\mathcal{H}')} |\bar\varphi\rangle\langle\bar\varphi|\nu(d\varphi d\psi),\end{aligned} \tag{20}$$

where the bar denotes complex conjugation in the basis$\{|i\rangle\}_{i=1}^{+\infty}$. By this equality for arbitrary Borel $B \subseteq \mathfrak{P}(\mathcal{H}')$ the operator

$$M(B) = \sigma^{-1/2} \left[\int\limits_{\mathfrak{P}(\mathcal{H})} \int\limits_{B} |\bar{\varphi}\rangle\langle\bar{\varphi}| \nu(d\varphi d\psi) \right] \sigma^{-1/2}$$

can be defined as a bounded positive operator on $\mathcal{H}$ such that $M(B) \leq M(\mathcal{X}) = I_{\mathcal{H}}$. It is easy to see that $M(d\psi)$ is a POVM on $\mathcal{X} = \mathfrak{P}(\mathcal{H}')$.

Consider the entanglement-breaking channel

$$\hat{\Phi}(\rho) = \int\limits_{\mathfrak{P}(\mathcal{H}')} |\psi\rangle\langle\psi| \mu_\rho(d\psi),$$

where μ_ρ is the Borel probability measure defined by $\mu_\rho(B) = \mathrm{Tr}\rho M(B), B \subseteq \mathcal{X}$. To prove that $\Phi(\rho) = \hat{\Phi}(\rho)$, it is sufficient to show that

$$\hat{\Phi}(|i\rangle\langle j|) = \Phi(|i\rangle\langle j|)$$

for all i, j. But

$$\hat{\Phi}(|i\rangle\langle j|) = \int\limits_{\mathfrak{P}(\mathcal{H}')} |\psi\rangle\langle\psi| \langle j|M(d\psi)|i\rangle$$
$$= \lambda_i^{-1/2} \lambda_j^{-1/2} \int\limits_{\mathfrak{P}(\mathcal{H})} \int\limits_{\mathfrak{P}(\mathcal{H}')} \langle i|\varphi\rangle\langle\varphi|j\rangle |\psi\rangle\langle\psi| \nu(d\varphi d\psi) = \Phi(e_{ij}),$$

where

$$e_{ij} = \lambda_i^{-1/2} \lambda_j^{-1/2} \mathrm{Tr}_{\mathcal{H}} (|j\rangle\langle i| \otimes I) |\Omega\rangle\langle\Omega| = |i\rangle\langle j|. \quad \square$$

The representation (15) by no means unique. Let us compare it with the Stinespring-Kraus representation [7] which is also not unique. However it is well known that that a (normal) completely positive map in a separable Hilbert space always has a discrete Stinespring-Kraus representation even if it is initially given as the corresponding integral. It is natural therefore to ask if for arbitrary entanglement-breaking channel there is a discrete representation of the form (15). As shown in [10], the answer is negative, which is closely related of existence of separable but countably indecomposable states.

6. Continuity Properties of the Entropy

As it is well known, the properties of the entropy for infinite and finite dimensional Hilbert spaces differ quite substantially: in the latter case the entropy is bounded continuous function on $\mathfrak{S}(\mathcal{H})$, while in the former it is discontinuous (lower semicontinuous) at every point, and infinite "most everywhere" in the sense that the set of states with finite entropy is a first category subset of $\mathfrak{S}(\mathcal{H})$ [19].

In what follows $\log$ denotes the function on $[0, +\infty)$, which coincides with the usual logarithm on $(0, +\infty)$ and vanishes at zero. If A is a positive finite rank operator in $\mathcal{H}$, then the entropy is defined as

$$H(A) = \mathrm{Tr}A\,(I \log \mathrm{Tr}A - \log A)\,, \tag{21}$$

where I is the unit operator in $\mathcal{H}$. If A, B two such operators then the relative entropy is defined as

$$H(A\,;B) = \mathrm{Tr}(A \log A - A \log B + B - A) \tag{22}$$

provided $\mathrm{ran}A \subseteq \mathrm{ran}B$, and $H(A\,;B) = +\infty$ otherwise (throughout this paper ran denotes the closure of the range of an operator in $\mathcal{H}$).

These definitions can be extended to arbitrary positive A,$B \in \mathfrak{T}(\mathcal{H})$ with the help of the following lemma [14]:

Lemma 2. *Let $\{P_n\}$ be an arbitrary sequence of finite dimensional projectors monotonously increasing to the unit operator I. The sequences $\{H(P_nAP_n)\}$, $\{H(P_nAP_n; P_nBP_n)\}$ are monotonously increasing and have the limits in the range $[0, +\infty]$ independent of the choice of the sequence $\{P_n\}$.*

We thus define the entropy and the relative entropy for density operators $\rho, \rho\,'$ as

$$H(\rho) = \lim_{n\to+\infty} H(P_n\rho P_n); \qquad H(\rho\,;\rho') = \lim_{n\to+\infty} H(P_n\rho P_n; P_n\rho' P_n).$$

Note that for fixed n $H(P_n\rho P_n), H(P_n\rho P_n; P_n\rho' P_n)$ are continuous functions of ρ, ρ'. This implies

Corollary. *Quantum entropy and relative entropy are lower semicontinuous in the following sense: let $\{\rho_n\}$ (resp. $\{\rho'_n\}$) be a sequence of density operators in $\mathcal{H}$, converging to a density operator ρ (resp. ρ'), then*

$$H(\rho) \le \lim\inf_{n\to\infty} H(\rho_n),$$

$$H(\rho;\rho') \le \lim\inf_{n\to\infty} H(\rho_n;\rho'_n).$$

Moreover, the finite dimensional approximation implies that the quantum entropy is concave function of ρ while the relative entropy is jointly convex function of $\rho, \rho\,'$ [14].

Let E be an $\mathfrak{H}$-operator. Denote

$$\beta(E) = \inf\{\beta > 0 : \mathrm{Tr}\exp(-\beta E) < \infty\}\,,$$

and let $\beta(E) = +\infty$, if $\mathrm{Tr}\exp(-\beta E) = +\infty$ for all $\beta > 0$.

Proposition 2. *Let $\mathcal{A} = \{\rho : \mathrm{Tr}\rho E \le h\}$. If $\beta(E) < \infty$, then the entropy $H(\rho)$ is bounded on $\mathcal{A}$; if $\beta(E) = 0$, then it is continuous on $\mathcal{A}$.*

Proof. (cf. [19]). Consider the Gibbs state $\rho_\beta = (\mathrm{Tr}\exp(-\beta E))^{-1}\exp(-\beta E)$. For arbitrary ρ in $\mathcal{A}$ we have

$$H(\rho;\rho_\beta) = -H(\rho) + \beta\mathrm{Tr}\rho E + \log \mathrm{Tr}\exp(-\beta E), \tag{23}$$

hence $H(\rho) \le \beta\mathrm{Tr}\rho E + \log \mathrm{Tr}\exp(-\beta E)$ which is finite for $\beta > \beta(E)$.

Let $\{\rho_n\}$ be an arbitrary sequence of states in $\mathcal{A}$ converging to the state ρ. By using (23) and lower semicontinuity of the relative entropy we obtain

$$\limsup_{n\to\infty} H(\rho_n) = H(\rho) + H(\rho;\rho_\beta) - \liminf_{n\to\infty} H(\rho_n;\rho_\beta) + \limsup_{n\to\infty} \beta \mathrm{Tr}\rho_n E - \beta \mathrm{Tr}\rho E \leq H(\rho) + \beta h.$$

By tending β in the above inequality to zero we establish the upper semicontinuity of the restriction of the function $H(\rho)$ to the set $\mathcal{A}$. Hence the restriction of the function $H(\rho)$ to the set $\mathcal{A}$ is continuous. □

7. The χ-Capacity of Constrained Channels

Consider a channel $\Phi : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H}')$, and let $\mathcal{A}$ be a subset of $\mathfrak{S}(\mathcal{H})$. We impose the constraint on (finite) input ensembles $\pi = \{\pi_i, \rho_i\}$ of the channel, by the requirement $\bar{\rho}(\pi) \equiv \sum_i \pi_i \rho_i \in \mathcal{A}$. The channel Φ with this constraint is called the $\mathcal{A}$ *-constrained* channel. The χ*-capacity* of the $\mathcal{A}$ -constrained channel Φ is defined as

$$\bar{C}(\Phi;\mathcal{A}) = \sup_{\bar{\rho}(\pi)\in\mathcal{A}} \chi_\Phi(\{\pi_i,\rho_i\}), \tag{24}$$

where

$$\chi_\Phi(\{\pi_i,\rho_i\}) = \sum_i \pi_i H(\Phi(\rho_i);\Phi(\bar{\rho})). \tag{25}$$

Importance of this quantity lies in the fact that it is closely related to the capacity of the channel for transmission of classical information (under the corresponding constraint onto the input states), namely

$$C(\Phi;\mathcal{A}) = \lim_{n\to\infty} \bar{C}(\Phi;\mathcal{A}),$$

see, e. g. [8]. If the χ-capacity is additive for given channel, then it coincides with the classical capacity. For example, this holds for the entanglement-breaking channels, as shown in [18] in finite dimensional case and generalized in [17] to infinite dimensions.

The subset of $\mathcal{P}_{\mathfrak{S}(\mathcal{H})}$, consisting of all ensembles π with the average $\bar{\rho}(\pi)$ in a subset $\mathcal{A} \subseteq \mathfrak{S}(\mathcal{H})$, will be denoted $\mathcal{P}_\mathcal{A}$. Consider the subset $\mathcal{P}_{\{\rho\}}$ consisting of all ensembles with the fixed average ρ. The proof of the following statement can be found in [9].

Lemma 3. *Collection of all finite ensembles from $\mathcal{P}_{\{\rho\}}$ is weakly dense in $\mathcal{P}_{\{\rho\}}$.*

Lemma 2 implies, in particular, that the nonnegative function $\rho \mapsto H(\Phi(\rho);\Phi(\bar{\rho}(\pi)))$ is measurable on $\mathfrak{S}(\mathcal{H})$. Hence the functional

$$\chi_\Phi(\pi) = \int\limits_{\mathfrak{S}(\mathcal{H})} H(\Phi(\rho);\Phi(\bar{\rho}(\pi)))\pi(d\rho)$$

is well defined on the set $\mathcal{P}_{\mathfrak{S}(\mathcal{H})}$ (with the range $[0;+\infty]$).

We are going to show that

$$\bar{C}(\Phi;\mathcal{A}) = \sup_{\pi\in\mathcal{P}_{\mathcal{A}}} \chi_\Phi(\pi). \tag{26}$$

This will follow from Lemma 3 and the following

Proposition 3. *The functional $\chi_\Phi(\pi)$ is lower semicontinuous on $\mathcal{P}_{\mathfrak{S}(\mathcal{H})}$. If $H(\Phi(\bar\rho(\pi)) < \infty$, then*

$$\chi_\Phi(\pi) = H(\Phi(\bar\rho(\pi))) - \int\limits_{\mathfrak{S}(\mathcal{H})} H(\Phi(\rho))\pi(d\rho). \tag{27}$$

Proof. Let $\{P_n\}$ be an arbitrary sequence of finite dimensional projectors monotonously increasing to the unit operator I. We show first that the functionals

$$\chi^n_\Phi(\pi) = \int\limits_{\mathfrak{S}(\mathcal{H})} H(P_n\Phi(\rho)P_n; P_n\Phi(\bar\rho(\pi))P_n)\pi(d\rho)$$

are continuous.

We have

$$\mathrm{ran}(P_n\Phi(\rho)P_n) \subseteq \mathrm{ran}(P_n\Phi(\bar\rho(\pi))P_n)$$

for $\pi-$almost all ρ. Indeed, closure of the range is orthogonal complement to the null subspace of a Hermitian operator, and for null subspaces the opposite inclusion holds obviously. It follows that

$$H(P_n\Phi(\rho)P_n; P_n\Phi(\bar\rho(\pi))P_n) = \mathrm{Tr}((P_n\Phi(\rho)P_n)\log(P_n\Phi(\rho)P_n)$$

$$-(P_n\Phi(\rho)P_n)\log(P_n\Phi(\bar\rho(\pi))P_n) + P_n\Phi(\bar\rho(\pi))P_n - P_n\Phi(\rho)P_n)$$

for $\pi-$almost all ρ. By using (21) we have

$$\chi^n_\Phi(\pi) = -\int\limits_{\mathfrak{S}(\mathcal{H})} H(P_n\Phi(\rho)P_n)\pi(d\rho) + \int\limits_{\mathfrak{S}(\mathcal{H})} \mathrm{Tr}(P_n\Phi(\rho)P_n)\log\mathrm{Tr}(P_n\Phi(\rho)P_n)\pi(d\rho)$$

$$-\int\limits_{\mathfrak{S}(\mathcal{H})} \mathrm{Tr}(P_n\Phi(\rho)P_n)\log(P_n\Phi(\bar\rho(\pi))P_n)\pi(d\rho).$$

The last term can be transformed in the following way

$$-\int\limits_{\mathfrak{S}(\mathcal{H})} \mathrm{Tr}(P_n\Phi(\rho)P_n)\log(P_n\Phi(\bar\rho(\pi))P_n)\pi(d\rho)$$

$$= -\mathrm{Tr}\int\limits_{\mathfrak{S}(\mathcal{H})} (P_n\Phi(\rho)P_n)\log(P_n\Phi(\bar\rho(\pi))P_n)\pi(d\rho)$$

$$= H(P_n\Phi(\bar\rho(\pi))P_n) - \mathrm{Tr}(P_n\Phi(\bar\rho(\pi))P_n)\log\mathrm{Tr}(P_n\Phi(\bar\rho(\pi))P_n).$$

Hence

$$\chi^n_\Phi(\pi) = -\int\limits_{\mathfrak{S}(\mathcal{H})} H(P_n\Phi(\rho)P_n)\pi(d\rho) + \int\limits_{\mathfrak{S}(\mathcal{H})} \mathrm{Tr}(P_n\Phi(\rho)P_n)\log \mathrm{Tr}(P_n\Phi(\rho)P_n)\pi(d\rho)$$

$$+H(P_n\Phi(\bar\rho(\pi))P_n) - \mathrm{Tr}(P_n\Phi(\bar\rho(\pi))P_n)\log \mathrm{Tr}(P_n\Phi(\bar\rho(\pi))P_n). \tag{28}$$

Continuity and boundedness of the quantum entropy in the finite dimensional case and similar properties of the function $\rho \mapsto \mathrm{Tr}(P_n\Phi(\rho)P_n)\log \mathrm{Tr}(P_n\Phi(\rho)P_n)$ imply continuity of the functionals $\chi^n_\Phi(\pi)$.

By the monotonous convergence theorem the sequence of functionals $\chi^n_\Phi(\pi)$ is nondecreasing and pointwise converges to $\chi_\Phi(\pi)$. Hence the functional $\chi_\Phi(\pi)$ is lower semicontinuous.

To prove (27) note that lemma 2 implies

$$\lim_{n\to+\infty} H(P_n\Phi(\bar\rho(\pi))P_n) = H(\Phi(\bar\rho(\pi)))$$

and

$$\lim_{n\to+\infty} \int\limits_{\mathfrak{S}(\mathcal{H})} H(P_n\Phi(\rho)P_n)\pi(d\rho) = \int\limits_{\mathfrak{S}(\mathcal{H})} H(\Phi(\rho))\pi(d\rho)$$

due to the monotonous convergence theorem. For every ρ the sequence $\{\mathrm{Tr}(P_n\Phi(\rho)P_n)\}$ is in $[0,1]$ and converges to 1, therefore $\lim_{n\to+\infty}\mathrm{Tr}(P_n\Phi(\rho)P_n)\log\mathrm{Tr}(P_n\Phi(\rho)P_n) = 0$, in particular the last term in (28) tends to 0. Since $|x\log x| < 1$ for all $x \in (0,1]$, the second term also tends to 0 by dominated convergence theorem, so passing to the limit $n\to\infty$ in (28) gives (27).□

Corollary 2. *The χ-capacity is given by the expression (26).*

Proof. The definition (24) is a similar expression in which the supremum is over all measures in $\mathcal{P}_\mathcal{A}$ with finite support. By lemma 3 we can approximate arbitrary measure π in $\mathcal{P}_\mathcal{A}$ by a sequence $\{\pi_n\}$ of measures in $\mathcal{P}_\mathcal{A}$ with finite support. By proposition 3, $\liminf_{n\to+\infty}\chi_\Phi(\pi_n) \geq \chi_\Phi(\pi)$. It follows that the supremum over all measures in $\mathcal{P}_\mathcal{A}$ coincides with the supremum over all measures in $\mathcal{P}_\mathcal{A}$ with finite support.□

Theorem 6. *Let $\mathcal{A}$ be a compact set. If the restriction of the output entropy $H(\Phi(\rho))$ to the set $\mathcal{A}$ is continuous then there exists an optimal ensemble π^* in $\mathcal{P}_\mathcal{A}$ such that* $\mathrm{supp}\pi^* \subseteq \mathfrak{P}(\mathcal{H})$ *and*

$$\bar C(\Phi;\mathcal{A}) = \chi_\Phi(\pi^*) = \int\limits_{\mathfrak{S}(\mathcal{H})} H(\Phi(\rho);\Phi(\bar\rho(\pi^*)))\pi^*(d\rho).$$

Proof. We will show first that the function

$$\pi \mapsto \int\limits_{\mathfrak{S}(\mathcal{H})} H(\Phi(\rho))\pi(d\rho)$$

is well defined and lower semicontinuous on the set $\mathcal{P}_\mathcal{A}$.

By lemma 2 the function $H(\Phi(\rho))$ is a pointwise limit of noncreasing sequences of functions

$$H(P_n\Phi(\rho)P_n) = \mathrm{Tr}\left((P_n\Phi(\rho)P_n)\left(I\log\mathrm{Tr}(P_n\Phi(\rho)P_n) - \log(P_n\Phi(\rho)P_n)\right)\right),$$

which are continuous and bounded on $\mathfrak{S}(\mathcal{H})$. Hence the function $H(\Phi(\rho))$ is measurable and the m.c.-theorem implies

$$\int\limits_{\mathfrak{S}(\mathcal{H})} H(\Phi(\rho))\pi(d\rho) = \lim_{n\to\infty}\int\limits_{\mathfrak{S}(\mathcal{H})} H(P_n\Phi(\rho)P_n)\pi(d\rho).$$

The sequence of continuous functionals

$$\pi \mapsto \int\limits_{\mathfrak{S}(\mathcal{H})} H(P_n\Phi(\rho)P_n)\pi(d\rho)$$

is nondecreasing. Hence its pointwise limit is lower semicontinuous.

By the assumption the restriction of the function $H(\Phi(\rho))$ to the set $\mathcal{A}$ is continuous hence $H(\Phi(\bar\rho(\pi))) < +\infty$ for any $\pi \in \mathcal{P}_\mathcal{A}$, and the relation (27) holds. Therefore the restriction of the functional $\chi_\Phi(\pi)$ to the set $\mathcal{P}_\mathcal{A}$ is upper semicontinuous, and by proposition 1 it is continuous. By theorem 4 the set $\mathcal{P}_\mathcal{A}$ is compact, hence $\chi_\Phi(\pi)$ attains its maximum on some $\pi_0 \in \mathcal{P}_\mathcal{A}$. Let us show that the maximum is attained on $\pi \in \mathcal{P}_\mathcal{A}$ supported by $\mathfrak{P}(\mathcal{H}) = \mathrm{ext}\mathfrak{S}(\mathcal{H})$.

Let $\{\pi_i^n, \rho_i^n\}$ be a sequence of finite ensembles with the averages $\bar\rho^n \in \mathcal{A}$ such that

$$\lim_{n\to+\infty}\chi_\Phi(\{\pi_i^n,\rho_i^n\}) = \bar C(\Phi;\mathcal{A}). \tag{29}$$

Since $\mathcal{A}$ is compact, we can assume that the averages $\bar\rho^n$ converge to a state $\bar\rho_0 \in \mathcal{A}$. Decomposing each state of the ensemble $\{\pi_i^n, \rho_i^n\}$ into a countable convex combination of pure states we obtain the new sequence $\{\hat\pi_j^n, \hat\rho_j^n\}$ of ensembles consisting of countable number of pure states with the same averages $\bar\rho^n$. Let $\hat\pi^n$ be the sequence of measures ascribing probability $\hat\pi_j^n$ to the one-point set $\{\hat\rho_j^n\}$ for each j. Convexity and lower semicontinuity of the relative entropy imply

$$\chi_\Phi(\hat\pi_n) = \sum_{j=1}^{\infty}\hat\pi_j^n H(\Phi(\hat\rho_j^n);\Phi(\bar\rho^n)) \geq \sum_i \pi_i^n H(\Phi(\rho_i^n);\Phi(\bar\rho^n)) = \chi_\Phi(\{\pi_i^n,\rho_i^n\}). \tag{30}$$

By construction $\mathrm{supp}\hat\pi^n \subseteq \mathrm{ext}\mathfrak{S}(\mathcal{H})$ for each n. Since $\hat\pi^n$ are in $\mathcal{P}_\mathcal{A}$ which is compact by theorem 4, there exists a subsequence $\hat\pi^{n_k}$, converging to some $\pi^* \in \mathcal{P}_\mathcal{A}$. Since the set $\mathrm{ext}\mathfrak{S}(\mathcal{H})$ of all pure states is closed subset of $\mathfrak{S}(\mathcal{H})$, we have $\mathrm{supp}\pi^* \subseteq \mathrm{ext}\mathfrak{S}(\mathcal{H})$ due to theorem 6.1 in [15]. By continuity of the average, $\bar\rho(\pi^*) = \bar\rho_0$. This, (29) and (30) imply

$$\bar C(\Phi;\mathcal{A}) = \lim_{k\to\infty}\chi_\Phi(\{\pi_i^{n_k},\rho_i^{n_k}\}) \leq \lim_{k\to\infty}\chi_\Phi(\hat\pi_{n_k}) = \chi_\Phi(\pi^*),$$

which means that π^* is an optimal ensemble for the $\mathcal{A}$ -constrained channel Φ.□

Corollary 3. *Let E, E' be $\mathfrak{H}$-operators in $\mathcal{H}'$ such that $\beta(E') = 0$ and $\mathrm{Tr}\rho E \leq h$* implies $\mathrm{Tr}\,\Phi(\rho)E' \leq h'$, *where $h, h\prime$ are some constants. Let*

$$\mathcal{A} = \{\rho : \mathrm{Tr}\rho E \leq h\}\,. \tag{31}$$

Then there exists an optimal ensemble for the $\mathcal{A}$ -constrained channel Φ supported by the set of pure states.

Proof. The set $\mathcal{A}$ is compact by theorem 3. Proposition 2 implies that the restriction of the entropy to the set $\mathcal{A}' = \Phi(\mathcal{A})$ is continuous, hence taking into account continuity of the map Φ we obtain that the restriction of the output entropy $H(\Phi(\rho))$ to the set $\mathcal{A}$ is continuous, which implies validity of the conditions of the theorem 6. □

8. Gaussian Channels

Bosonic system with finite number s of degrees of freedom, such as collection of s modes of the radiation field, is mathematically described in terms of symplectic space. Let $\mathbf{R}^s$ be Euclidean space of vectors α, β with the inner product $\langle\beta, \alpha\rangle = \beta^T\alpha$. The symplectic space $Z = \mathbf{R}^s \oplus \mathbf{R}^s$ consists of vectors $z = (\alpha, \beta)$ and is equipped with the form

$$\Delta(z, z') = \langle\beta', \alpha\rangle - \langle\beta, \alpha'\rangle = (z')^T\Delta z,$$

where

$$\Delta = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

is the skew-symmetric commutation matrix, see e. g. [11].

Let $\mathcal{H}$ be the Hilbert space $L^2(\mathbf{R}^s)$ of complex square-integrable functions of the real variables $\xi = (\xi_j)_{j=\overline{1,s}}$. Consider the two strongly continuous groups of unitary operators acting on $\psi \in \mathcal{H}$ according to the formulas

$$U_\alpha\psi(\xi) = \psi(\xi + \alpha); \quad V_\beta\psi(\xi) = \exp(i\langle\beta, \xi\rangle)\psi(\xi). \tag{32}$$

The operators $U_\alpha, V_\beta,$ form a couple of unitary representations of $\mathbf{R}^s$, satisfying the Weyl canonical commutation relations

$$U_\alpha V_\beta = \exp i\langle\beta, \alpha\rangle V_\beta U_\alpha. \tag{33}$$

Defining the operators $W_z = \exp\frac{i}{2}\langle\beta, \alpha\rangle V_\beta U_\alpha$ for $z = (\alpha, \beta)$, one has a projective unitary representation $z \to W_z$ of the additive group $Z = \mathbf{R}^s \oplus \mathbf{R}^s$, namely

$$W_z W_{z'} = \exp\frac{i}{2}\Delta(z, z')W_{z+z'}.$$

This is implies the Weyl-Segal commutation relations

$$W_{z'}W_z = \exp i\Delta(z, z')W_z W_{z'}. \tag{34}$$

Stone's theorem implies $W_z = \exp iR^T z$, where $R = (p_j; q_j)_{j=\overline{1,s}}$ is the vector of canonical observables of the system, satisfying the Heisenberg commutation relations on

a dense domain in $\mathcal{H}$. Starting with the unitary Weyl operators rather than unbounded canonical observables not only allows to avoid irrelevant pathologies but is also more convenient operationally.

Characteristic function of a quantum state ρ in $\mathcal{H}$ is defined as

$$\phi(z) = \mathrm{Tr}\rho W(z); \quad z \in Z.$$

The state ρ is called *Gaussian*, if its characteristic function $\phi(z) = \mathrm{Tr}\rho W(z)$ has the form

$$\phi(z) = \exp\left(i\, m^T z - \frac{1}{2} z^T \alpha z\right), \tag{35}$$

where m is the *mean vector* and α is the *correlation matrix*, so that

$$m = \mathrm{Tr}\rho R\ ;\ \alpha - \frac{i}{2}\Delta = \mathrm{Tr}(R-m)\rho(R-m)^T. \tag{36}$$

A channel Φ in $\mathcal{H}$ is called *Gaussian* [11] if its dual acts on the Weyl operators as

$$\Phi^*[W_z] = f(z) W_{K^T z}, \tag{37}$$

where $f(z)$ is exponential of a quadratic form and K is a linear operator in Z. By inspection of the characteristic functions, it follows that Φ maps Gaussian states into Gaussian states, which may be taken for alternative definition. However, most transparently the Gaussian channels are described in terms of the canonical transformations involving environment.

Let $\mathcal{H}$ be the Hilbert space of the system, and $\mathcal{H}_e$ – that of the environment, and let R, R_e be the vectors of canonical observables in $\mathcal{H}, \mathcal{H}_e$. Let ρ and ρ_e be the Gaussian density operators in $\mathcal{H}$ and $\mathcal{H}_e$ with the correlation matrices α and α_e, respectively. For simplicity of notations we take zero mean vectors in what follows. Consider the linear transformation

$$R' = KR + K_e R_e \tag{38}$$

where K, K_e are real matrices (to simplify notations we write R, R_e instead of $R \otimes I_e, I \otimes R_e$ etc.) satisfying

$$\Delta = K\Delta K^T + K_e \Delta K_e^T.$$

This condition guarantees that R' are again the canonical variables satisfying the commutation relations. In terms of the Weyl operators the transformation (8) is

$$W'(z) = W(K^T z) \otimes W_e(K_e^T z).$$

Then for a fixed state ρ_e of the environment, the channel transformation taking the input state ρ to the output state ρ' is described in terms of characteristic functions as

$$\phi'(z) = \phi(K^T z)\phi_e\left(K_e^T z\right) \tag{39}$$

which is just the form (37) with

$$f(z) = \phi_e\left(K_e^T z\right) = \mathrm{Tr}\rho_e W_e(K_e^T z) \tag{40}$$

Thus the factor f is simply expressed in terms of the characteristic function of the initial state of the environment. Conversely, starting from (37) and using Stinespring-Kraus representations, one can find dynamical equations (8) for the channel.

The correlation matrix of the output observables R' is computed via (36) with $m = 0$, namely

$$\alpha' - \frac{i}{2}\Delta' = \mathrm{Tr}R'\rho' R'^T,$$

where $\rho' = \rho \otimes \rho_e$. Then using (8), we obtain

$$\alpha' = K\alpha K^T + K_e\alpha_e K_e^T. \tag{41}$$

The conditions of corollary 3 are fulfilled for any Gaussian channel with the power constraint of the form (31) where $E = R^T\epsilon R$ is the many-mode oscillator Hamiltonian with arbitrary nondegenerate energy matrix ϵ and R are the canonical variables of the system. Then E is $\mathfrak{H}$-operator with $\beta(E) = 0$, and by theorem 3 the constraint set $\mathcal{A} = \{\rho : \mathrm{Tr}\rho E \leq h\}$ is compact. In terms of correlation matrix α of the input state ρ, the constraint (31) takes the form $\mathrm{Sp}\alpha\epsilon \leq h$, where Sp denotes trace of matrix. Then, taking

$$E' = cR^T R, \quad (c > 0),$$

we again have $\beta(E') = 0$. Using (41) we obtain $\mathrm{Tr}\Phi(\rho)E' = c\mathrm{Sp}(\alpha K^T K + \alpha_e K_e^T K_e)$. Now we can always choose a positive c such that $cK^T K \leq \epsilon$, hence $\rho \in \mathcal{A}$ implies $\mathrm{Tr}\Phi(\rho)E' \leq h + c\mathrm{Sp}(\alpha_e K_e^T K_e) \equiv h'$.

In classical information theory the Shannon capacity of Gaussian channels with quadratic power constraints is known to be achieved on Gaussian inputs. It is natural to ask if a similar result holds for quantum Gaussian channels.

Conjecture. *For Gaussian channel with the power constraint an optimal generalized ensemble is given by a Gaussian measure supported by the set of pure Gaussian states with arbitrary mean and a fixed correlation matrix.*

One instance of genuinely quantum channel for which this holds is the pure loss channel [5] which however has the very special feature that its minimal output entropy is zero and the minimum is achieved on coherent states. The Gaussian ensemble of the coherent states then indeed maximizes the χ-capacity in accordance with formulas for more general attenuation/amplification channel [11]. For one mode $(s = 1)$ this channel is characterized by the following action on the Weyl operators

$$\Phi^*[W_z] = W_{kz}\exp\left[-\frac{\hbar}{2}\left(|k^2 - 1|/2 + N_e\right)|z|^2\right], \tag{42}$$

where k is attenuation/amplification coefficient, N_e is the mean photon number of the environment. Let the input state ρ of the channel be the thermal Gaussian (see [11]) with mean photon number N and the entropy

$$H(\rho) = g(N) \equiv (N+1)\log(N+1) - N\log N. \tag{43}$$

Then the output state $\Phi[\rho]$ is again thermal Gaussian with N replaced by $N' = k^2N + N_0'$, where $N_0' = \max\{0, (k^2-1)\} + N_e$ is the output mean photon number corresponding to the input vacuum state. Thus the output entropy is equal to

$$H(\Phi[\rho]) = g(N'). \tag{44}$$

Define the *constrained* $\chi-$capacity by the formula (24), where the supremum is taken over all ensembles π, satisfying the power constraint $\mathrm{Tr}\bar{\rho}(\pi)a^\dagger a \leq N$. It was conjectured in [11] that this supremum is realized by the ensemble of coherent states $\rho_z = |z\rangle\langle z|$ with the Gaussian probability density $p(z) = (2\pi N)^{-1}\exp\left(-|z|^2/2N\right)$, giving the value

$$\bar{C}(\Phi) = g\left(N'\right) - g\left(N_0'\right).$$

In the case of pure attenuation the equation (8) in terms of annihilation operators takes the form

$$a' = ka + \sqrt{1-k^2}a_e,$$

where additionally $N_e = 0$. As observed in [5], the above conjecture can be verified in this case, moreover, the additivity

$$\bar{C}(\Phi^{\otimes n}) = n\bar{C}(\Phi) \tag{45}$$

holds, and hence

$$C(\Phi) = \bar{C}(\Phi) = g(k^2N). \tag{46}$$

Indeed, for this channel $N_0' = 0, N' = k^2N$ and $H(\Phi[\rho]) = g(k^2N)$ if the input is the elementary Gaussian state with the mean photon number N. Taking $N = 0$ gives $H(\Phi[\rho_z]) = 0$ for arbitrary coherent state ρ_z, and thus the minimal output entropy vanishes implying its additivity. The relations (45) and (46) then follow from observation that $\bar{C}(\Phi^{\otimes n})$ is just the maximal output entropy under the power constraint. Similar considerations allow to prove additivity and compute $C(\Phi)$ for pure loss channel with arbitrary number s of the modes, see [5].

Acknowledgments

The author gratefully acknowledges the support and hospitality of the organizers of the meeting.

References

[1] E. Alfsen, *Compact convex sets and boundary integrals*, Springer, 1971.

[2] O. Bratteli, D.W. Robinson, *Operators algebras and quantum statistical mechanics*, Springer Verlag, New York-Heidelberg-Berlin, vol.I, 1979.

[3] E. B. Davies, Quantum stochastic processes-I, *Comm. Math. Phys.* **15** (1969), 277-304.

[4] G. F. Dell'Antonio, On the limits of sequences of states, *Commun. Pure Appl. Math.* **20** (1967), 413-430.

[5] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J.H. Shapiro and H. P. H. Yuen, Classical capacity of the lossy bosonic channel: exact solution, quant-ph/0308012.

[6] J. Eisert, C. Simon, M. B. Plenio, On the quantification of the entanglement in infinite-dimensional quantum systems, quant-ph/0112064.

[7] A. S. Holevo, *Statistical structure of quantum theory*, Springer 2001.

[8] A. S. Holevo, Classical capacities of quantum channels with constrained inputs, *Probability Theory and Applications* **48** (2003), 359-374; quant-ph/0211170.

[9] A. S. Holevo, M. E. Shirokov, Continuous ensembles and the χ-capacity of infinite dimensional channels, *Probability Theory and Applications* **50** (2005), 98-114; quant-ph/0408176.

[10] A. S. Holevo, M. E. Shirokov, R. F. Werner, Separability and entanglement-breaking in infinite dimensions, quant-ph/0504204.

[11] A. S. Holevo, R. F. Werner, Evaluating capacities of Bosonic Gaussian channels, *Phys. Rev*. **A63** (2001), 032312; quant-ph/9912067.

[12] M. Horodecki, P.W. Shor, M.B. Ruskai, General Entanglement Breaking Channels, *Rev. Math. Phys.* **15** (2003), 629-641; quant-ph/0302031

[13] A. D. Joffe, W. M. Tikhomirov, *Theory of extremum problems*, Academic Press, New York, 1979.

[14] G. Lindblad, Expectation and entropy inequalities for finite quantum systems, *Comm. Math. Phys.* **39** (1974), 111-119.

[15] K. Parthasarathy, *Probability measures on metric spaces,* Academic Press, New York and London, 1967.

[16] T. A. Sarymsakov, *Introduction to quantum probability theory,* FAN, Tashkent, 1985 (In Russian).

[17] M. E. Shirokov, The Holevo capacity of infinite dimensional channels, quant-ph/0408009.

[18] P. W. Shor, Additivity of the classical capacity of entanglement-breaking quantum channel, quant-ph/0201149.

[19] A. Wehrl, General properties of entropy, *Rev. Mod. Phys.* **50** (1978), 221-250.

[20] R. F. Werner, M. M. Wolf, Bound entangled Gaussian states, e-print quant-ph/0009118.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Introduction to Relativistic Quantum Information

Daniel Terno [1]
Perimeter Institute for Theoretical Physics,
31 Caroline St., Waterloo, Ontario, Canada N2L 2Y5

Abstract. I discuss the role that relativistic considerations play in quantum information processing. First I describe how the causality requirements limit possible multi-partite measurements. Then the Lorentz transformations of quantum states are introduced, and their implications on physical qubits are described. This is used to describe relativistic effects in communication and entanglement.

To the memory of Asher Peres, teacher and friend

1. Introduction

Information and physics are closely and fascinatingly intertwined. Their relations become even more interesting when we leave a non-relativistic quantum mechanics for more exiting venues. My notes are planned as a guided tour for the first steps along that road, with open questions and more involved mergers left to the remarks and to the last section.

I start from a brief introduction to causality restrictions on the distributed measurements: the limitations that are imposed by final propagation velocity of the physical interactions. It is followed by the relativistic transformations of the states of massive particles and photons, from which we can deduce what happens to qubits which are realized as the discrete degrees of freedom. Building on this, I discuss the distinguishability of quantum signals, and briefly touch communication channels and the bipartite entanglement.

I do not follow a historical order or give all of the original references. A review [1] is used as the standard reference on quantum information and relativity. The results of the "usual" quantum information are given without any reference: all of them can be found in at least one of the sources [2,3,4]. Finally, a word about units: $\hbar = c = 1$ are always assumed.

2. Causality and Distributed Measurements

Here I present the causality constraints on quantum measurements. For simplicity, measurements are considered to be point-like interventions. First recall the standard descrip-

[1] e-mail: dterno@perimetrinstitute.ca

tion of the measurement and the induced state transformation. Consider a system in the state ρ that is subject to measurement that is described by a positive operator-valued measure (POVM) $\{E_\mu\}$. The probability of the outcome μ is

$$p_\mu = \mathrm{tr}\, E_\mu \rho, \tag{1}$$

while the state transformation is given by some completely positive evolution

$$\rho \to \rho'_\mu = \sum_m A_{\mu m} \rho A^\dagger_{\mu m}/p_\mu, \qquad \sum_m A^\dagger_{\mu m} A_{\mu m} = E_\mu. \tag{2}$$

If the outcome is left unknown, the update rule is

$$\rho \to \rho = \sum_{\mu m} A_{\mu m} \rho A^\dagger_{\mu m}. \tag{3}$$

Now consider a bipartite state ρ_{AB}. The operations of Alice and Bob are given by the operators $A_{\mu m}$ and $B_{\nu n}$, respectively. It is easy to see that if these operators commute,

$$[A_{\mu m}, B_{\nu n}] = 0, \tag{4}$$

then the observation statistics of Bob is independent of Alice's results and vice versa. Indeed, the probability that Bob gets a result ν, irrespective of what Alice found, is

$$p_\nu = \sum_\mu \mathrm{tr}\left(\sum_{m,n} B_{\nu n}\, A_{\mu m}\, \rho\, A^\dagger_{\mu m}\, B^\dagger_{\nu n} \right). \tag{5}$$

Now make use of Eq. (4) to exchange the positions of $A_{\mu m}$ and $B_{\nu n}$, and likewise those of $A^\dagger_{\mu m}$ and $B^\dagger_{\nu n}$, and then we move $A_{\mu m}$ from the first position to the last one in the product of operators in the traced parenthesis. Since the elements of a POVM satisfy $\sum_\mu E_\mu = \mathbb{1}$, Eq. (5) reduces to

$$p_\nu = \mathrm{tr}\left(\sum_n B_{\nu n}\, \rho\, B^\dagger_{\nu n} \right), \tag{6}$$

whence all the expressions involving Alice's operators $A_{\mu m}$ have totally disappeared. The statistics of Bob's result are not affected at all by what Alice may simultaneously do somewhere else. This proves that Eq. (4) indeed is a sufficient condition for no instantaneous information transfer. In particular, the local operations $A \otimes \mathbb{1}_B$ and $\mathbb{1}_A \otimes B$ are of this form.

Note that any classical communication between distant observers can be considered a kind of long range interaction. The propagation of signals is, of course, bounded by the velocity of light. As a result, there exists a partial time ordering of the various interventions in an experiment, which defines the notions earlier and later. The input parameters of an intervention are deterministic (or possibly stochastic) functions of the parameters of earlier interventions, but not of the stochastic outcomes resulting from later or mutually spacelike interventions [1].

Even these apparently simple notions lead to non-trivial results. Consider a separable bipartite superoperator T,

$$T(\rho) = \sum_k M_k \rho M_k^\dagger, \qquad M_k = A_k \otimes B_k, \tag{7}$$

where the operators A_k represent operations of Alice and B_k those of Bob. Not all such superoperators can be implemented by local transformations and classical communication (LOCC) [7]. This is the foundation of the "non-locality without entanglement".

A classification of bipartite state transformations was introduced in [8]. It consists of the following categories. There are *localizable* operations that can be implemented locally by Alice and Bob, possibly with the help of prearranged ancillas, but without classical comunication. Ideally, local operations are instantaneous, and the whole process can be viewed as performed at a definite time. A final classical output of such distributed intervention will be obtained at some point of the (joint) causal future of Alice's and Bob's interventions. For *semilocalizable* operations, the requirement of no communication is relaxed and one-way classical communication is possible. It is obvious that any tensor-product operation $T_\mathrm{A} \otimes T_\mathrm{B}$ is localizable, but it is not a necessary condition. For example the Bell measurements, which distinguishes between the four standard bipartite entangled qubit states,

$$|\Psi^\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle|1\rangle \pm |1\rangle|0\rangle), \qquad |\Phi^\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle|0\rangle \pm |1\rangle|1\rangle), \tag{8}$$

are localizable.

Other classes of bipartite operators are defined as follows: Bob performs a local operation T_B just before the global operation T. If no local operation of Alice can reveal any information about T_B, i.e., Bob cannot signal to Alice, the operation T is *semicausal*. If the operation is semicausal in both directions, it is *causal*. In many cases it is easier to prove causality than localizability (see Remark 3). There is a necessary and sufficient condition for the semicausality (and therefore, the causality) of operations [8].

These definitions of causal and localizable operators appear equivalent. It is easily proved that localizable operators are causal. It was shown that semicausal operators are always semilocalizable [9]. However, there are causal operations that are not localizable [8].

It is curious that while a complete Bell measurement is causal, the two-outcome incomplete Bell measurement is not. Indeed, consider a two-outcome PVM

$$E_1 = |\Phi^+\rangle\langle\Phi^+|, \qquad E_2 = \mathbb{1} - E_1. \tag{9}$$

If the initial state is $|01\rangle_\mathrm{AB}$, then the outcome that is associated with E_2 always occurs and Alice's reduced density matrix after the measurement is $\rho_\mathrm{A} = |0\rangle\langle 0|$. On the other hand, if before the joint measurement Bob performs a unitary operation that transforms the state into $|00\rangle_\mathrm{AB}$, then the two outcomes are equiprobable, the resulting states after the measurement are maximally entangled, and Alice's reduced density matrix is $\rho_\mathrm{A} = \frac{1}{2}\mathbb{1}$. A simple calculation shows that after this incomplete Bell measurement two input states $|00\rangle_\mathrm{AB}$ and $|01\rangle_\mathrm{AB}$ are distinguished by Alice with a probability of 0.75.

Here is another example of a semicausal and semilocalizable measurement which can be executed with one-way classical communication from Alice to Bob. Consider a PVM measurement, whose complete orthogonal projectors are

$$|0\rangle \otimes |0\rangle, \quad |0\rangle \otimes |1\rangle, \quad |1\rangle|+\rangle, \quad |1\rangle \otimes |-\rangle, \tag{10}$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. The Kraus matrices are

$$A_{\mu j} = E_\mu \delta_{j0}, \tag{11}$$

From the properties of complete orthogonal measurements [8], it follows that this operation cannot be performed without Alice talking to Bob. A protocol to realize this measurement is the following. Alice measures her qubit in the basis $\{|0\rangle, |1\rangle\}$, and tells her result to Bob. If Alice's outcome was $|0\rangle$, Bob measures his qubit in the basis $\{|0\rangle, |1\rangle\}$, and if it was $|1\rangle$, in the basis $\{|+\rangle, |-\rangle\}$.

If one allows for more complicated conditional state evolution [10], then more measurements are localizable. In particular, consider a *verification* measurement, i.e., the measurement yields a μ-th result with certainty, if the state prior to the classical interventions was given by $\rho = E_\mu$, but without making any specific demand on the resulting state ρ'_μ.

It is possible to realize a verification measurements by means of a shared entangled ancilla and Bell-type measurements by one of the parties [11]. Verification measurement of Eq. (10) can illustrate this construction. In addition to the state to be tested, Alice and Bob share a Bell state $|\Psi^-\rangle$. They do not have to coordinate their moves. Alice and Bob perform tasks independently and convey their results to a common center, where a final decision is made.

The procedure is based on the teleportation identity

$$|\Psi\rangle_1|\Psi^-\rangle_{23} = \tfrac{1}{2}\left(|\Psi^-\rangle_{12}|\Psi\rangle_3 + |\Psi^+\rangle_{12}|\tilde{\Psi}^{(z)}\rangle_3 + |\Phi^-\rangle_{12}|\tilde{\Psi}^{(x)}\rangle_3 + |\Phi^+\rangle_{12}|\tilde{\Psi}^{(y)}\rangle_3\right), \tag{12}$$

where $|\tilde{\Psi}^{(z)}\rangle$ means the state $|\Psi\rangle$ rotated by π around the z-axis, etc. The first step of this measurement corresponds to the first step of a teleportation of a state of the spin from B (Bob's site) to A (Alice's site). Bob and Alice do not perform the full teleportation (which requires a classical communication between them). Instead, Bob performs only the Bell measurement at his site which leads to one of the branches of the superposition in the rhs of Eq. (12).

The second step of the verification measurement is taken by Alice. Instead of completing the teleportation protocol, she measures the spin of her particle in the z direction. According to whether that spin is up or down, she measures the spin of her ancilla in the z or x direction, respectively. This completes the measurement and it only remains to combine the local outcomes to get the result of the nonlocal measurement [11]. This method can be extended to arbitrary Hilbert space dimensions.

Remarks

1. Measurements in quantum field theory are discussed in [1,5,6].

2. An algebraic field theory approach to statistical independence and to related topics is presented in [12].

3. To check the causality of an operation T whose outcomes are the states $\rho_\mu = T_\mu(\rho)/p_\mu$ with probabilities $p_\mu = \mathrm{tr}\, T_\mu(\rho)$, $\sum_\mu p_\mu = 1$ it is enough to consider the corresponding superoperator

$$T'(\rho) := \sum_\mu T_\mu(\rho) \tag{13}$$

Indeed, assume that Bob's action prior to the global operation lead to one of the two different states ρ_1 and ρ_2. Then the states $T'(\rho_1)$ and $T'(\rho_2)$ are distinguishable if and only if some of the pairs of states $T_\mu(\rho_1)/p_{\mu 1}$ and $T_\mu(\rho_2)/p_{\mu 2}$ are distinguishable. Such probabilistic distinguishability shows that the operation T is not semicausal.

3. Absence of the superluminal communication makes possible to evade the theorems on the impossibility of a bit commitment. In particular the protocol RBC2 allows a bit commitment to be indefinitely maintained with unconditionally security against all classical attacks, and at least for some finite amount of time against quantum attacks [13,14].

4. In these notes I am not going to deal with the relativistic localization POVM. Their properties (and difficulties in their construction) can be found in [1]. An exhaustive survey of the spatial localization of photons is given in [15]. Here we only note in passing that if $E(\mathcal{O})$ is an operator that corresponds to the detection of an event in a spacetime region $\mathcal{O}$, since they are not thought to be implemented by physical operations confined to that spacetime area, the condition $[E(\mathcal{O}_1), E(\mathcal{O}_2)] = 0$ is not required [16,17].

3. Quantum Lorentz Transformations

There is no elementary particle that is called "qubit". Qubits are realized by particular degrees of freedom of more or less complicated systems. To decide how qubits transform (e.g., under Lorentz transformations) it may be necessary to consider again the entire system. In the following our qubit will be either a spin of a massive particle or a polarization of a photon. A quantum Lorentz transformation connects the description of a quantum state $|\Psi\rangle$ in two reference frames that are connected by a Lorentz transformation Λ (i.e., their coordinate axes are rotated with respect to each other and the frames have a fixed relative velocity). Then $|\Psi'\rangle = U(\Lambda)|\Psi\rangle$, and the unitary $U(\Lambda)$ is represented on Fig. 1 below. The purpose of this section is to explain the elements of this quantum circuit.

From the mathematical point of view the single-particle states belong to some irreducible representation of the Poincaré group. An introductory discussion of these representations and their relations with states and quantum fields may be found, e.g., in [18,19]. Within each particular irreducible representation there are six commuting operators. The eigenvalues of two of them are invariants that label the representation by defining the mass m and the intrinsic spin j. The basis states are labelled by three components of the momentum $\mathbf{p}$ and the spin operator Σ_3. Hence a generic state is given by

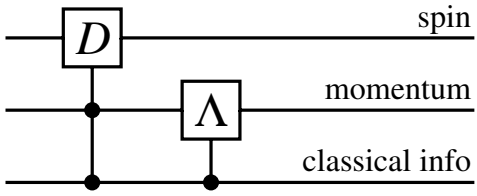


Figure 1. Relativistic state transformation as a quantum circuit: the gate D which represents the matrix $D_{\xi\sigma}[W(\Lambda,p)]$ is controlled by both the classical information and the momentum p, which is itself subject to the classical information Λ.

$$|\Psi\rangle = \sum_{\sigma}\int d\mu(p)\psi_{\sigma}(p)|p,\sigma\rangle. \tag{14}$$

In this formula $d\mu(p)$ is the Lorentz-invariant measure,

$$d\mu(p) = \frac{1}{(2\pi)^3}\frac{d^3\mathbf{p}}{2E(\mathbf{p})}, \tag{15}$$

where the energy $E(\mathbf{p}) = p^0 = \sqrt{\mathbf{p}^2 + m^2}$. The improper momentum and spin eigenstates are δ-normalized,

$$\langle p,\sigma|q,\sigma'\rangle = (2\pi)^3(2E(\mathbf{p}))\delta^{(3)}(\mathbf{p}-\mathbf{q})\delta_{\sigma\sigma'}, \tag{16}$$

and are complete on the one-particle space, which is $\mathcal{H} = \mathbb{C}^{2j+1}\otimes L^2(\mathbb{R}^3, d\mu(p))$ for spin-j fields.

To find the transformation law we have to be more concrete about the spin operator. The operator $\Sigma_3(p)$ is a function of the generators of the Poincaré group. One popular option is helicity, $\Sigma_3 = \mathbf{J}\cdot\mathbf{P}/|\mathbf{P}|$, which is applicable for both massive and massless particles. For massive particles we use the z-component of the rest-frame (or Wigner spin, that we now describe in the next section.

3.1. Massive particles

The construction involve picking a reference 4-momentum k, which for massive particles is taken to be $k_R = (m,\mathbf{0})$. The Wigner spin $\mathbf{S}(p)$ is defined to coincide with the non-relativistic spin $\mathbf{S}$ in particle's rest frame. The state of a particle at rest is labelled $|k_R,\sigma\rangle$,

$$\mathbf{S}^2|k_R,\sigma\rangle = j(j+1)|k_R,\sigma\rangle, \qquad \mathbf{S}_3|k_R,\sigma\rangle = \sigma|k_R,\sigma\rangle. \tag{17}$$

The spin states of arbitrary momenta are defined as follows. The standard rotation-free boost that brings k_R to an arbitrary momentum p, $p^\mu = L(p)^\mu{}_\nu k^\nu$ is given by

$$L(p) = \begin{pmatrix} \frac{E}{m} & \frac{p_1}{m} & \frac{p_2}{m} & \frac{p_3}{m} \\ \frac{p_1}{m} & 1+\frac{p_1^2}{m(m+E)} & \frac{p_1p_2}{m(m+E)} & \frac{p_1p_3}{m(m+E)} \\ \frac{p_2}{m} & \frac{p_2p_1}{m(m+E)} & 1+\frac{p_2^2}{m(m+E)} & \frac{p_2p_3}{m(m+E)} \\ \frac{p_3}{m} & \frac{p_3p_1}{m(m+E)} & \frac{p_3p_2}{m(m+E)} & 1+\frac{p_3^2}{m(m+E)} \end{pmatrix}. \tag{18}$$

The Wigner spin $\mathbf{S}(p)$ and the one-particle basis states are defined by

$$|p,\sigma\rangle \equiv U[L(p)]|k_R,\sigma\rangle, \qquad \mathbf{S}_3(p)|p,\sigma\rangle = \sigma|p,\sigma\rangle. \tag{19}$$

In deriving the transformation rules we begin with the momentum eigenstates. Using the group representation property and Eqs. (19) the transformation is written as

$$U(\Lambda) = U[L(\Lambda p)]U[L^{-1}(\Lambda p)\Lambda.L(p)]U[L^{-1}(p)] \tag{20}$$

The element of the Lorentz group

$$W(\Lambda,p) \equiv L^{-1}(\Lambda p)\Lambda L(p), \tag{21}$$

leaves k_R invariant, $k_R = Wk_R$. Hence it belongs to the stability subgroup (or Wigner little group) of k_R. For $k_R = (m,\mathbf{0})$ it is a rotation. Pressing on

$$U(\Lambda)|p,\sigma\rangle = U[L(\Lambda p)]U[W(\Lambda,p)]|k_R,\sigma\rangle, \tag{22}$$

and as a result,

$$U(\Lambda)|p,\sigma\rangle = \sum_\xi D_{\xi\sigma}[W(\Lambda,p)]|\Lambda p,\xi\rangle, \tag{23}$$

where $D_{\xi\sigma}$ are the matrix elements of the representation of the Wigner rotation $W(\Lambda,p)$.

We consider only spin-$\frac{1}{2}$ particles, so $\sigma = \pm\frac{1}{2}$. Any 2×2 unitary matrix can be written as $\hat{D} = \exp(-i\omega\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})$, where ω is a rotation angle and $\hat{\mathbf{n}}$ is a rotation axis that corresponds to $W(\Lambda,p)$.

The wave functions transform according to $\psi'_\xi(q) = \langle\xi,q|U(\Lambda)|\Psi\rangle$ so the same state in the Lorentz-transformed frame is

$$|\Psi'\rangle = U(\Lambda)|\Psi\rangle = \sum_{\sigma,\xi}\int_{-\infty}^{\infty} D_{\sigma\xi}[W(\Lambda,\Lambda^{-1}p)]\psi_\xi(\Lambda^{-1}p)|\sigma,p\rangle d\mu(p). \tag{24}$$

For pure rotation $\mathcal{R}$ the three-dimensional (more exactly, 3D block of 4D matrix; here and in the following we use the same letter for a 4D and 3D matrix for $\mathcal{R}\in$SO(3)) Wigner rotation matrix is the rotation itself,

$$W(\mathcal{R},p) = \mathcal{R}, \qquad \forall p = (p^0,\mathbf{p}). \tag{25}$$

As a result, the action of Wigner spin operators on $\mathcal{H}_1$ is given by than halves of Pauli matrices that are tensored with the identity of L^2.

3.2. Photons

The single-photon states are labelled by momentum $\mathbf{p}$ (the 4-momentum vector is null, $E = p^0 = |\mathbf{p}|$) and helicity $\sigma_\mathbf{p} = \pm 1$, so the state with a definite momentum is given by $\sum_{\sigma=\pm1}\alpha_\sigma|p,\sigma_\mathbf{p}\rangle$, where $|\alpha_+|^2 + |\alpha_-|^2 = 1$. Polarization states are also labelled by 3-vectors $\boldsymbol{\epsilon}^\sigma_\mathbf{p}$, $\mathbf{p}\cdot\boldsymbol{\epsilon}^\sigma_\mathbf{p} = 0$, that correspond to the two senses of polarization of

classical electromagnetic waves. An alternative labelling of the same state, therefore, is $\sum_{\sigma=\pm1}\alpha_\sigma|p,\epsilon^\sigma_{\mathbf{p}}\rangle$.

Helicity is invariant under proper Lorentz transformation, but the basis states acquire phases. The little group element $W(\Lambda,p)=L^{-1}(\Lambda p)\Lambda L(p)$ is defined with respect to the standard four-momentum $k_R=(1,0,0,1)$. The standard Lorentz transformation is

$$L(p)=R(\hat{\mathbf{p}})B_z(u), \tag{26}$$

where $B_z(u)$ is a pure boost along the z-axis with a velocity u that takes k_R to $(|\mathbf{p}|,0,0,|\mathbf{p}|)$ and $R(\hat{\mathbf{p}})$ is the standard rotation that carries the z-axis into the direction of the unit vector $\hat{\mathbf{p}}$. If $\hat{\mathbf{p}}$ has polar and azimuthal angles θ and ϕ, the standard rotation $R(\hat{\mathbf{p}})$ is accomplished by a rotation by θ around the y-axis, that is followed by a rotation by ϕ around the z-axis. Hence,

$$R(\hat{\mathbf{p}})=\begin{pmatrix}\cos\theta\cos\phi & -\sin\phi & \cos\phi\sin\theta\\ \cos\theta\sin\phi & \cos\phi & \sin\phi\sin\theta\\ -\sin\theta & 0 & \cos\theta\end{pmatrix}, \tag{27}$$

(here only the non-trivial 3D block is shown).

An arbitrary little group element for a massless particle is decomposed according to

$$W(\Lambda,p)=S(\beta,\gamma)R_z(\xi), \tag{28}$$

where the elements $S(\beta,\gamma)$ form a subgroup that is isomorphic to the translations of the Euclidean plane and $R_z(\xi)$ is a rotation around the z-axis. We are interested only in the angle ξ, since β and γ do not correspond to the physical degrees of freedom. However, they are important for gauge transformations. Finally, the little group elements are represented by

$$D_{\sigma'\sigma}=\exp(i\xi\sigma)\delta_{\sigma'\sigma}. \tag{29}$$

It is worthwhile to derive more explicit expressions for ξ. I begin with rotations, $\Lambda=\mathcal{R}$. Since rotations form a subgroup of a Lorentz group, $R^{-1}(\mathcal{R}\hat{\mathbf{p}})\mathcal{R}R(\hat{\mathbf{p}})$ is a rotation that leaves $\hat{\mathbf{z}}$ invariant and thus is of the form $R_z(\omega)$ for some ω. A boost in (t,z) plane and a rotation around z-axis commute, $[R_z,B_z]=0$, so

$$W(\mathcal{R},p)=R^{-1}(\mathcal{R}\hat{\mathbf{p}})\mathcal{R}R(\hat{\mathbf{p}})=R_z(\xi). \tag{30}$$

Any rotation can be described by two angles that give a direction of the axis and the third angle that gives the amount of rotation around that axis. If $\mathcal{R}\mathbf{p}=\mathbf{q}$, we decompose the rotation matrix as

$$\mathcal{R}=R_{\hat{\mathbf{q}}}(\omega)R(\hat{\mathbf{q}})R^{-1}(\hat{\mathbf{p}}), \tag{31}$$

where $R_{\hat{\mathbf{q}}}(\omega)$ characterizes a rotation around $\hat{\mathbf{q}}$, and $R(\hat{\mathbf{q}})$ and $R(\hat{\mathbf{p}})$ are the standard rotations that carry the z-axis to $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$, respectively. Using Eq. (30) we find that $S=\mathbb{1}$ and the two rotations are of the same conjugacy class,

$$R_z(\xi) = R^{-1}(\mathcal{R}\hat{\mathbf{p}})R_{\mathcal{R}\hat{\mathbf{p}}}(\omega)R(\mathcal{R}\hat{\mathbf{p}}), \tag{32}$$

so we conclude that $\xi = \omega$.

A practical description of polarization states is given by spatial vectors that correspond to the classical polarization directions. Taking again k_R as the reference momentum, two basis vectors of linear polarization are $\boldsymbol{\epsilon}^1_{k_R} = (1,0,0)$ and $\boldsymbol{\epsilon}^2_{k_R} = (0,1,0)$, while to the right and left circular polarizations correspond $\boldsymbol{\epsilon}^{\pm}_{k_R} = (\boldsymbol{\epsilon}^1_{k_R} \pm i\boldsymbol{\epsilon}^2_{k_R})/\sqrt{2}$.

Phases of the states obtained by the standard Lorentz transformations $L(p)$ are set to 1. Since the standard boost $B_z(u)$ leaves the four-vector $(0, \boldsymbol{\epsilon}^{\pm}_{k_R})$ invariant, we define a polarization basis for any $\mathbf{p}$ as

$$\boldsymbol{\epsilon}^{\pm}_{\mathbf{p}} = \boldsymbol{\epsilon}^{\pm}_{\hat{\mathbf{p}}} \equiv R(\hat{\mathbf{p}})\boldsymbol{\epsilon}^{\pm}_{k_R}, \tag{33}$$

while the transformation of polarization vectors under an arbitrary rotation $\mathcal{R}$ is given by the rotation itself. To see the agreement between transformations of spatial vectors and states, consider a generic state with a momentum $\mathbf{p}$. Its polarization is described by the polarization vector $\boldsymbol{\alpha}(\mathbf{p}) = \alpha_+\boldsymbol{\epsilon}^+_{\mathbf{p}} + \alpha_-\boldsymbol{\epsilon}^-_{\mathbf{p}}$, or by the state vector $\alpha_+|\mathbf{p},+\rangle + \alpha_-|\mathbf{p},-\rangle$. Using Eq. (33) we see that the transformation of $\boldsymbol{\alpha}(\mathbf{p})$ is given by

$$\mathcal{R}\boldsymbol{\alpha}(\mathbf{p}) = R_{\mathcal{R}\hat{\mathbf{p}}}(\omega)R(\mathcal{R}\hat{\mathbf{p}})R^{-1}(\hat{\mathbf{p}})\boldsymbol{\alpha}(\mathbf{p}) = R_{\mathcal{R}\hat{\mathbf{p}}}(\omega)R(\mathcal{R}\hat{\mathbf{p}})\boldsymbol{\alpha}(k_R) = R_{\mathcal{R}\hat{\mathbf{p}}}(\omega)\boldsymbol{\alpha}(\mathcal{R}\hat{\mathbf{p}}). \tag{34}$$

If $\mathbf{q} = \mathcal{R}\mathbf{p}$ the transformation results in $\alpha_+e^{i\omega}\boldsymbol{\epsilon}^+_{\mathbf{q}} + \alpha_-e^{-i\omega}\boldsymbol{\epsilon}^-_{\mathbf{q}}$, and since $\omega = \xi$, it is equivalent to the state transformation

$$U(\mathcal{R})(\alpha_+|p,+\rangle + \alpha_-|p,-\rangle) = \alpha_+e^{i\xi}|q,+\rangle + \alpha_-e^{-i\xi}|q,-\rangle. \tag{35}$$

For a general Lorentz transformations the triad $(\boldsymbol{\epsilon}^1_{\mathbf{p}}, \boldsymbol{\epsilon}^2_{\mathbf{p}}, \hat{\mathbf{p}})$ is rigidly rotated, but in a more complicated fashion. To obtain the phase for a general Lorentz transformation, we decompose the latter into two rotations and a standard boost B_z along the z-axis:

$$\Lambda = \mathcal{R}_2B_z(u)\mathcal{R}_1. \tag{36}$$

It can be shown that B_z alone does not lead to a phase rotation. Therefore,

$$\xi = \omega_1 + \omega_2, \tag{37}$$

where both ω_1 and ω_2 are due to the rotations and are given by Eq. (31). Note that although $B_z(u)$ alone does not lead to a phase rotation, it can affect the value of ω_2, since it indirectly appears in the definition of $\mathcal{R}_2$.

Remarks

1. A comprehensive discussion of the Poincaré group in physics can be found in [20,21]. Useful expressions for Wigner rotations and their applications for massive particles are given in [22,23,24].

2. In this transformation I do not assume any additional normalization factors. A condition of unitarity is $UU^\dagger = U^\dagger U = \mathbb{1}$, but there also other conventions in the literature.

3. A double infinity of the positive energy solutions of the Dirac equation (functions $u_p^{(1/2)}$ and $u_p^{(-1/2)}$) spans an improper basis of this space. There is a one-to-one correspondence between Wigner and Dirac wave functions. Basis vectors of Wigner and Dirac Hilbert spaces are in the one-to-one correspondence [21],

$$u_p^{(1/2)} \Leftrightarrow |\tfrac{1}{2}, p\rangle, \qquad u_p^{(-1/2)} \Leftrightarrow |\tfrac{1}{2}, p\rangle, \tag{38}$$

while the wave functions are related by

$$\Psi^\alpha(p) = \psi_{1/2}(p) u_p^{(1/2)\alpha} + \psi_{-1/2}(p) u_p^{(-1/2)\alpha} \tag{39}$$

$$2m\psi_\sigma(p) = \sum_{\alpha=1}^{4} \overline{u}_{\alpha p}^{(-\sigma)} \Psi^\alpha(p) \tag{40}$$

4. Another approach to the construction of the Wigner rotation $\hat{D}$ is based on the homomorphism between Lorentz group and $SL(2)$ [21].

5. When not restricted to a single-particle space the Wigner spin operator is given by

$$\mathbf{S} = \tfrac{1}{2} \sum_{\eta,\zeta} \boldsymbol{\sigma}_{\eta\zeta} \int d\mu(p) (\hat{a}_{\eta p}^\dagger \hat{a}_{\zeta p} + \hat{b}_{\eta p}^\dagger \hat{b}_{\sigma p}), \tag{41}$$

where $\hat{a}_{\eta p}^\dagger$ creates a mode with a momentum p and spin η along the z-axis, etc. A comparison of different spin operators can be found in [25].

6. If one works with the 4-vectors, then in the helicity gauge the polarization vector is given by $\epsilon_p = (0, \boldsymbol{\epsilon}_{\mathbf{p}})$. A formal connection between helicity states and polarization vectors is made by first observing that three spin-1 basis states can be constructed from the components of a symmetric spinor of rank 2. Unitary transformations of this spinor that are induced by $\mathcal{R}$ are in one-to-one correspondence with transformations by $\mathcal{R}$ of certain linear combinations of a spatial vector. In particular, transformations of the helicity ± 1 states induced by rotations are equivalent to the rotations of $(\boldsymbol{\epsilon}_{k_S}^1 \pm i\boldsymbol{\epsilon}_{k_S}^2)/\sqrt{2}$ (the z-axis is the initial quantization direction). While $p_\mu \epsilon_p^\mu = 0$ gauge condition is Lorentz-invariant, the spatial orthogonality is not. The role of gauge transformations in preserving the helicity gauge and some useful expressions for the phase that photons acquire can be found in [26,27,28]

4. Implications of Quantum Lorentz Transformations

4.1. Reduced Density Matrices

In a relativistic system whatever is outside the past light cone of the observer is unknown to him, but also cannot affect his system, therefore does not lead to decoherence (here, I assume that no particle emitted by from the outside the past cone penetrates into the future cone). Since different observers have different past light cones, by tracing out they

exclude from their descriptions different parts of spacetime. Therefore any transformation law between them must tacitly assume that the part excluded by one observer is irrelevant to the system of another.

Another consequence of relativity is that there is a hierarchy of dynamical variables: *primary variables* have relativistic transformation laws that depend only on the Lorentz transformation matrix Λ that acts on the spacetime coordinates. For example, momentum components are primary variables. On the other hand, *secondary variables* such as spin and polarization have transformation laws that depend not only on Λ, but also on the momentum of the particle. As a consequence, the reduced density matrix for secondary variables, which may be well defined in any coordinate system, has no transformation law relating its values in different Lorentz frames.

Moreover, an unambiguous definition of the reduced density matrix is possible only if the secondary degrees of freedom are unconstrained, and photons are the simplest example when this definition fails. In the absence of a general prescription, a case-by-case treatment is required. I describe a particular construction, valid with respect to a certain class of tests.

4.2. Massive Particles

For a massive qubit the usual definition of quantum entropy has no invariant meaning. The reason is that under a Lorentz boost, the spin undergoes a Wigner rotation, that as shown on Fig. 1 is controlled both by the classical data and the corresponding momentum. Even if the initial state is a direct product of a function of momentum and a function of spin, the transformed state is not a direct product. Spin and momentum become entangled.

Let us define a reduced density matrix,

$$\rho = \int d\mu(p)\psi(p)\psi^{\dagger}(p). \tag{42}$$

It gives statistical predictions for the results of measurements of spin components by an ideal apparatus which is not affected by the momentum of the particle. Note that I tacitly assumed that the relevant observable is the Wigner spin. The spin entropy is

$$S = -\mathrm{tr}\,(\rho \log \rho) = -\sum \lambda_j \log \lambda_j, \tag{43}$$

where λ_j are the eigenvalues of ρ.

As usual, ignoring some degrees of freedom leaves the others in a mixed state. What is not obvious is that in the present case the amount of mixing depends on the Lorentz frame used by the observer. Indeed consider another observer (Bob) who moves with a constant velocity with respect to Alice who prepared that state. In the Lorentz frame where Bob is at rest, the state is given by Eq. (24).

As an example, take a particle prepared by Alice to be

$$|\Psi\rangle = \chi \int \psi(p)|p\rangle d\mu(p), \qquad \chi = \begin{pmatrix} \zeta \\ \eta \end{pmatrix} \tag{44}$$

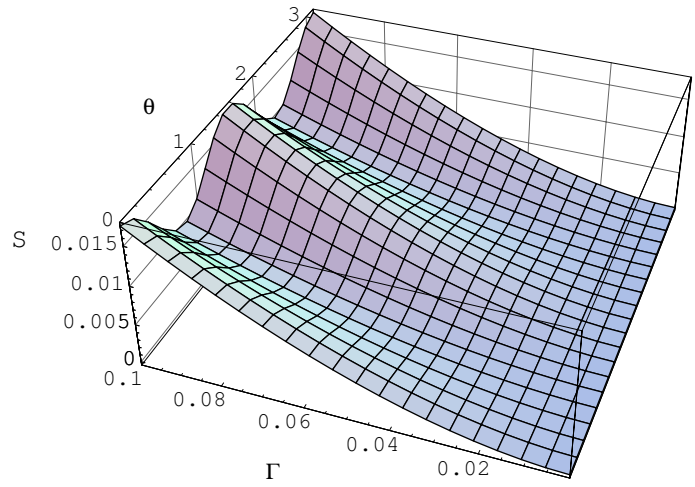


Figure 2. Dependence of the spin entropy S, in Bob's frame, on the values of the angle θ and a parameter $\Gamma = [1-(1-v^2)^{1/2}]\Delta/mv$.

where ψ is concentrated near zero momentum and has a characteristic spread Δ. Spin density matrices of all the states that are given by Eq. (44) are

$$\rho = \begin{pmatrix} |\zeta|^2 & \zeta\eta^* \\ \zeta^*\eta & |\eta|^2 \end{pmatrix}, \tag{45}$$

and are independent of the specific form of $\psi(p)$. To make calculations explicit (and simpler) I take the wave function to be Gaussian, $\psi(p) = N\exp(\mathbf{p}^2/2\Delta^2)$, where N is a normalization factor. Spin and momentum are not entangled, and the spin entropy is zero. When that particle is described in Bob's Lorentz frame, moving with velocity v at the angle θ with Alice's z-axis, a detailed calculation shows that the the spin entropy is positive [1]. This phenomenon is illustrated in Fig. 2. A relevant parameter, apart from the angle θ, is in the leading order in momentum spread,

$$\Gamma = \frac{\Delta}{m}\,\frac{1-\sqrt{1-v^2}}{v}, \tag{46}$$

where Δ is the momentum spread in Alice's frame. The entropy has no invariant meaning, because the reduced density matrix τ has no covariant transformation law, except in the limiting case of sharp momenta. Only the complete density matrix transforms covariantly.

I outline some of the steps in this derivation. First, we calculate the rotation parameters $(\omega, \hat{\mathbf{n}})$ of the orthogonal matrix $W(\Lambda, p)$ for a general momentum. The rotation axis and angle are given by

$$\hat{\mathbf{n}} = \hat{\mathbf{v}} \times \hat{\mathbf{p}}, \qquad \cos\theta = \hat{\mathbf{v}} \cdot \hat{\mathbf{p}}, \qquad 0 \leq \theta \leq \pi \tag{47}$$

where $\hat{\mathbf{v}}$ is boost's direction, while the leading order term for the angle is

$$\omega = \frac{1-\sqrt{1-v^2}}{v}\frac{p}{m}\sin\theta - \mathcal{O}\left(\frac{p^2}{m^2}\right) \tag{48}$$

Without a loss of generality we can make another simplification. We can choose our coordinate frame in such a way that both ζ and η are real. The matrix $D[W(\Lambda,\Lambda^{-1}p)]$ takes the form

$$D[W(\Lambda,p')] = \sigma_0\cos\frac{\omega}{2} - i\sin\frac{\omega}{2}(-\sin\phi\,\sigma_x + \cos\phi\,\sigma_y), \tag{49}$$

where (θ,ϕ) are the spherical angle of $\mathbf{p}'$ (to be consistent with Eq. (24) momentum in Alice frame carries a prime, $p' = \Lambda^{-1}p$). The reduced density matrix in Bob's frame is

$$\rho^B_{\sigma\xi} = \int d\mu(p) D_{\sigma\nu} D^*_{\xi\lambda}\psi_\nu(p')\psi^*_\lambda(p'). \tag{50}$$

The symmetry of $\psi(\Lambda^{-1}p)$ is cylindrical. Hence the partial trace is taken by performing a momentum integration in cylindrical coordinates. This simplification is a result of the spherical symmetry of the original ψ. The two remaining integrations are performed by first expanding in powers of p/Δ and taking Gaussian integrals. Finally,

$$\rho' = \begin{pmatrix} \zeta^2(1-\Gamma^2/4) + \eta^2\Gamma^2/4 & \zeta\eta^*(1-\Gamma^2/4) \\ \zeta^*\eta(1-\Gamma^2/4) & \zeta^2\Gamma^2/4 + \eta^2(1-\Gamma^2/4) \end{pmatrix}. \tag{51}$$

Fidelity can be used to estimate the difference between the two density matrices. It is defined as

$$f = \chi^\dagger\rho'\chi, \tag{52}$$

and it is easy to get an analytical result for this quantity. Set $\zeta = \cos\theta$ and $\eta = \sin\theta$. Then

$$f = 1 - \frac{\Gamma^2}{2}\left(3 + \frac{\cos 4\theta}{8}\right). \tag{53}$$

Consider now a pair of orthogonal states that were prepared by Alice, e.g. the above state with $\chi_1 = (1,0)$ and $\chi_2 = (0,1)$. How well can moving Bob distinguish them? I use the simplest criterion, namely the probability of error P_E, defined as follows: an observer receives a single copy of one of the two known states and performs any operation permitted by quantum theory in order to decide which state was supplied. The probability of a wrong answer for an optimal measurement is

$$P_E(\rho_1,\rho_2) = \tfrac{1}{2} - \tfrac{1}{4}\,\mathrm{tr}\sqrt{(\rho_1-\rho_2)^2}. \tag{54}$$

In Alice's frame $P_E = 0$. In Bob's frame the reduced density matrices are

$$\rho^B_1 = \begin{pmatrix} 1-\Gamma^2/4 & 0 \\ 0 & \Gamma^2/4 \end{pmatrix}, \qquad \rho^B_2 = \begin{pmatrix} \Gamma^2/4 & 0 \\ 0 & 1-\Gamma^2/4, \end{pmatrix} \tag{55}$$

respectively. Hence the probability of error is $P_E(\rho_1,\rho_2) = \Gamma^2/4$.

4.3. Photons

The relativistic effects in photons are essentially different from those for massive particles that were discussed above. This is because photons have only two linearly independent polarization states. As we know, polarization is a *secondary variable*: states that correspond to different momenta belong to distinct Hilbert spaces and cannot be superposed (an expression such as $|\boldsymbol{\epsilon}^{\pm}_{\mathbf{k}}\rangle + |\boldsymbol{\epsilon}^{\pm}_{\mathbf{q}}\rangle$ is meaningless if $\mathbf{k} \neq \mathbf{q}$). The complete basis $|p, \boldsymbol{\epsilon}^{\pm}_{\mathbf{p}}\rangle$ does not violate this superselection rule, owing to the orthogonality of the momentum basis. The reduced density matrix, according to the usual rules, should be

$$\rho = \int d\mu(p)|\psi(p)|^2 |p, \boldsymbol{\alpha}(\mathbf{p})\rangle\langle p, \boldsymbol{\alpha}(\mathbf{p})|. \tag{56}$$

However, since ξ in Eq. (29) depends on the photon's momentum even for ordinary rotations, this object will have no transformation law at all. It is still possible define an "effective" density matrix adapted to a specific method of measuring polarization [29,30]. I describe one such scheme.

The labelling of polarization states by Euclidean vectors $\boldsymbol{\epsilon}^{\pm}_{\mathbf{p}}$ suggests the use of a 3×3 matrix with entries labelled x, y and z. Classically, they correspond to different directions of the electric field. For example, a component ρ_{xx} would give the expectation values of operators representing the polarization in the x direction, seemingly irrespective of the particle's momentum.

To have a momentum-independent polarization is to admit longitudinal photons. Momentum-independent polarization states thus consist of physical (transverse) and unphysical (longitudinal) parts, the latter corresponding to a polarization vector $\boldsymbol{\epsilon}^{\ell} = \hat{\mathbf{p}}$. For example, a generalized polarization state along the x-axis is

$$|\hat{\mathbf{x}}\rangle = x_+(\mathbf{p})|\boldsymbol{\epsilon}^{+}_{\mathbf{p}}\rangle + x_-(\mathbf{p})|\boldsymbol{\epsilon}^{-}_{\mathbf{p}}\rangle + x_\ell(\mathbf{p})|\boldsymbol{\epsilon}^{\ell}_{\mathbf{p}}\rangle, \tag{57}$$

where $x_\pm(\mathbf{p}) = \hat{\mathbf{x}} \cdot \boldsymbol{\epsilon}^{\pm}_{\mathbf{p}}$, and $x_\ell(\mathbf{p}) = \hat{\mathbf{x}} \cdot \hat{\mathbf{p}} = \sin\theta\cos\phi$. It follows that $|x_+|^2 + |x_-|^2 + |x_\ell|^2 = 1$, and we thus define

$$\mathbf{e}_x(\mathbf{p}) = \frac{x_+(\mathbf{p})\boldsymbol{\epsilon}^{+}_{\mathbf{p}} + x_-(\mathbf{p})\boldsymbol{\epsilon}^{-}_{\mathbf{p}}}{\sqrt{x_+^2 + x_-^2}}, \tag{58}$$

as the polarization vector associated with the x direction. It follows from (57) that $\langle\hat{\mathbf{x}}|\hat{\mathbf{x}}\rangle = 1$ and $\langle\hat{\mathbf{x}}|\hat{\mathbf{y}}\rangle = \hat{\mathbf{x}} \cdot \hat{\mathbf{y}} = 0$, and likewise for the other directions, so that

$$|\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{x}}| + |\hat{\mathbf{y}}\rangle\langle\hat{\mathbf{y}}| + |\hat{\mathbf{z}}\rangle\langle\hat{\mathbf{z}}| = \mathbb{1}_p, \tag{59}$$

where $\mathbb{1}_p$ is the unit operator in momentum space.

To the direction $\hat{\mathbf{x}}$ there corresponds a projection operator

$$P_{xx} = |\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{x}}| \otimes \mathbb{1}_p = |\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{x}}| \otimes \int d\mu(k)|\mathbf{p}\rangle\langle\mathbf{p}|, \tag{60}$$

The action of P_{xx} on $|\Psi\rangle$ follows from Eq. (57) and $\langle\boldsymbol{\epsilon}^{\pm}_{\mathbf{p}}|\boldsymbol{\epsilon}^{\ell}_{\mathbf{p}}\rangle = 0$. Only the transverse part of $|\hat{\mathbf{x}}\rangle$ appears in the expectation value:

$$\langle\Psi|P_{xx}|\Psi\rangle = \int d\mu(p)|\psi(p)|^2|x_+(\mathbf{p})\alpha_+^*(\mathbf{p}) + x_-(\mathbf{p})\alpha_-^*(\mathbf{p})|^2. \tag{61}$$

It is convenient to write the transverse part of $|\hat{\mathbf{x}}\rangle$ as

$$|\mathbf{b}_x(\mathbf{p})\rangle \equiv (|\boldsymbol{\epsilon}_{\mathbf{p}}^+\rangle\langle\boldsymbol{\epsilon}_{\mathbf{p}}^+| + |\boldsymbol{\epsilon}_{\mathbf{p}}^-\rangle\langle\boldsymbol{\epsilon}_{\mathbf{p}}^-|)|\hat{\mathbf{x}}\rangle = x_+(\mathbf{p})|\boldsymbol{\epsilon}_{\mathbf{p}}^+\rangle + x_-(\mathbf{p})|\boldsymbol{\epsilon}_{\mathbf{p}}^-\rangle. \tag{62}$$

Likewise define $|\mathbf{b}_y(\mathbf{p})\rangle$ and $|\mathbf{b}_z(\mathbf{p})\rangle$. These three state vectors are neither of unit length nor mutually orthogonal.

Finally, a POVM element E_{xx} which is the physical part of P_{xx}, namely is equivalent to P_{xx} for physical states (without longitudinal photons) is

$$E_{xx} = \int d\mu(k)|p, \mathbf{b}_x(\mathbf{p})\rangle\langle p, \mathbf{b}_x(\mathbf{p})|, \tag{63}$$

and likewise for the other directions. The operators E_{xx}, E_{yy} and E_{zz} indeed form a POVM in the space of physical states, owing to Eq. (59).

To complete the construction of the density matrix, we introduce additional directions. Following the standard practice of state reconstruction, we consider $P_{x+z,x+z}$, $P_{x+iz,x+iz}$ and similar combinations. For example,

$$P_{x+z,x+z} = \tfrac{1}{2}(|\hat{\mathbf{x}}\rangle + |\hat{\mathbf{z}}\rangle)(\langle\hat{\mathbf{x}}| + \langle\hat{\mathbf{z}}|) \otimes \mathbb{1}_p. \tag{64}$$

The diagonal elements of the new polarization density matrix are defined as

$$\rho_{mm} = \langle\Psi|E_{mm}|\Psi\rangle, \qquad m = x, y, z, \tag{65}$$

and the off-diagonal elements are recovered by combinations such as

$$\rho_{xz} = \langle\Psi(|\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{z}}| \otimes \mathbb{1}_p)|\Psi\rangle = \langle\Psi|E_{x+z,x+z} - iE_{x-iz,x-iz} + (1-i)(E_{xx} - E_{zz})/2|\Psi\rangle. \tag{66}$$

Denote $|\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{z}}| \otimes \mathbb{1}_p$ as P_{xz}, and its "physical" part by E_{xz}. Then the effective polarization density matrix is

$$\rho_{mn} = \langle\Psi|E_{mn}|\Psi\rangle = \int d\mu(k)|f(\mathbf{p})|^2\langle\boldsymbol{\alpha}(\mathbf{p})|\mathbf{b}_m(\mathbf{p})\rangle\langle\mathbf{b}_n(\mathbf{p})|\boldsymbol{\alpha}(\mathbf{p})\rangle, \quad m, n, = x, y, z. \tag{67}$$

It is interesting to note that this derivation gives a direct physical meaning to the naive definition of a reduced density matrix,

$$\rho_{mn}^{\text{naive}} = \int d\mu(p)|\phi(p)|^2\boldsymbol{\alpha}_m(\mathbf{p})\boldsymbol{\alpha}_n^*(\mathbf{p}) = \rho_{mn} \tag{68}$$

It is possible to show that this POVM actually corresponds to a simple photodetection model [31].

The basis states $|p, \boldsymbol{\epsilon}_{\mathbf{p}}\rangle$ are direct products of momentum and polarization. Owing to the transversality requirement $\boldsymbol{\epsilon}_{\mathbf{p}} \cdot \mathbf{p} = 0$, they remain direct products under Lorentz transformations. All the other states have their polarization and momentum degrees of freedom entangled. As a result, if one is restricted to polarization measurements as de-

scribed by the above POVM, *there do not exist two orthogonal polarization states*. In general, any measurement procedure with finite momentum sensitivity will lead to the errors in identification, as demonstrated as follows

Let two states $|\Phi\rangle$ and $|\Psi\rangle$ be two orthogonal single-photon states. Their reduced polarization density matrices, ρ_Φ and ρ_Ψ, respectively, are calculated using Eq. (67). Since the states are entangled, the von Neumann entropies of the reduced density matrices, $S = -\mathrm{tr}(\rho \ln \rho)$, are positive. Therefore, both matrices are at least of rank two. Since the overall dimension is 3, it follows that $\mathrm{tr}(\rho_\Phi \rho_\Psi) > 0$ and these states are not perfectly distinguishable. An immediate corollary is that photon polarization states cannot be cloned perfectly, because the no-cloning theorem forbids exact copying of unknown non-orthogonal states.

In general, any measurement procedure with finite momentum sensitivity will lead to the errors in identification. First I present some general considerations and then illustrate them with a simple example. Let us take the z-axis to coincide with the average direction of propagation so that the mean photon momentum is $k_A\hat{\mathbf{z}}$. Typically, the spread in momentum is small, but not necessarily equal in all directions. Usually the intensity profile of laser beams has cylindrical symmetry, and we may assume that $\Delta_x \sim \Delta_y \sim \Delta_r$ where the index r means radial. We may also assume that $\Delta_r \gg \Delta_z$. We then have

$$f(p) \propto f_1[(p_z - k_A)/\Delta_z]\, f_2(p_r/\Delta_r). \tag{69}$$

We approximate

$$\theta \approx \tan\theta \equiv p_r/p_z \approx p_r/k_A. \tag{70}$$

In pictorial language, polarization planes for different momenta are tilted by angles up to $\sim \Delta_r/k_A$, so that we expect an error probability of the order Δ_r^2/k_A^2. In the density matrix ρ_{mn} all the elements of the form ρ_{mz} should vanish when $\Delta_r \to 0$. Moreover, if $\Delta_z \to 0$, the non-vanishing xy block goes to the usual (monochromatic) polarization density matrix,

$$\rho_{\mathrm{pure}} = \begin{pmatrix} |\alpha|^2 & \beta & 0 \\ \beta^* & 1-|\alpha|^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{71}$$

As an example, consider two states which, if the momentum spread could be ignored, would be $|k_A\hat{\mathbf{z}}, \boldsymbol{\epsilon}^{\pm}_{k_A\hat{\mathbf{z}}}\rangle$. To simplify the calculations we assume a Gaussian distribution:

$$f(p) = Ne^{-(p_z-k_A)^2/2\Delta_z^2} e^{-p_r^2/2\Delta_r^2}, \tag{72}$$

where N is a normalization factor and $\Delta_z, \Delta_r \ll k_A$. In general the spread in p_z may introduce an additional incoherence into density matrices, in addition to the effect caused by the transversal spread. However, when all momentum components carry the same helicity, this spread results in corrections of the higher order. In the example below we take the polarization components to be $\boldsymbol{\epsilon}^{\pm}_{\mathbf{p}} \equiv R(\hat{\mathbf{p}})\boldsymbol{\epsilon}^{\pm}_{k_S}$. That means we have to analyze the states

$$|\Psi_\pm\rangle = \int d\mu(p) f(p) |p, \boldsymbol{\epsilon}^\pm_{\mathbf{p}}\rangle, \tag{73}$$

where $f(p)$ is given above.

It is enough to expand $R(\hat{\mathbf{p}})$ up to second order in θ. The reduced density matrices are calculated by techniques similar to those for massive particles, using rotational symmetry around the z-axis and normalization requirements. At the leading order in $\Omega \equiv \Delta_r/k_A$

$$\rho_+ = \tfrac{1}{2}(1 - \tfrac{1}{2}\Omega^2) \begin{pmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \tfrac{1}{2}\Omega^2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{74}$$

and $\rho_- = \rho_+^*$. At the same level of precision,

$$P_E(\rho_+, \rho_-) = \Delta_r^2/4k_A^2. \tag{75}$$

It is interesting to note that the optimal strategy for distinguishing between these two states is a polarization measurement in the xy-plane. Then the effective 2×2 density matrices are perfectly distinguishable, but there is a probability $\Omega^2/2$ that no photon will be detected at all. The above result is valid due to the special form of the states that we had chosen. Potential errors in the upper 2×2 blocks are averaged out in the integration over ϕ. The effect becomes important when, e.g., a plain monochromatic wave undergoes a strong focusing. Then $\Omega \approx l/f$, where l is the aperture radius and f is the focal length [31].

Now let us turn to the distinguishability problem from the point of view of a moving observer, Bob. The probability of an error is still given by Eq. (54), but the parameters depend on the observer's motion. Assume again that Bob moves along the z-axis with a velocity v. To calculate Bob's reduced density matrix, we must transform the complete state, and then take a partial trace.

Reduced density matrices of $|\Psi_\pm\rangle$ in both frames are given by the expression

$$(\rho_\pm)^{A,B}_{mn} = \int d\mu(p) |f(p)^{A,B}|^2 \langle R(\hat{\mathbf{p}})\boldsymbol{\epsilon}^\pm_{k_S} | \mathbf{b}_m(\mathbf{p})\rangle \langle \mathbf{b}_n(\mathbf{p}) | R(\hat{\mathbf{p}})\boldsymbol{\epsilon}^\pm_{k_S}\rangle. \tag{76}$$

This is due to the following two reasons. First, $|\mathbf{b}_m(\mathbf{p})\rangle$ are defined by Eq. (62) in any frame, while pure boosts preserve the orientation of the coordinate axes in 3-space, and therefore do not affect the indices of ρ_{mn}. Second, phases acquired by polarization states cancel out, since we choose the states $|\mathbf{p}, \boldsymbol{\alpha}(\mathbf{p})\rangle$ to be the helicity eigenstates.

Calculation of Bob's density matrix is similar to the previous cases. The only frame dependent expression in (76) is $f^B(p) = f^A(\Lambda^{-1}p)$. A boost along the z-axis preserves k_r and ϕ. On the other hand,

$$k_z^B \approx k_A \sqrt{\frac{1-v}{1+v}}. \tag{77}$$

Since everything else in the integral remains the same, the effect of relative motion is given by a substitution

$$\Omega^B = \sqrt{\frac{1+v}{1-v}}\Omega^A, \qquad P_E^B = \frac{1+v}{1-v}P_E^A, \tag{78}$$

so Bob can distinguish the signals either better or worse than Alice [29].

Remarks

1. A modification of the spin operator [32] will allow for a momentum-independent transformation of the spin density matrix between two frames that are related by a fixed Lorentz transformation Λ_{12}. Its relation to our scheme is discussed in [33].

2. An additional motivation for introduction of effective polarization density matrices cames from the analysis of one-photon scattering [30].

3. I have discussed only discrete variables. To explore the relativistic effects with continuous variables [34] it is convenient to express the quantum Lorentz transformations in terms of mode creation and annihilation operators [35].

5. Communication channels

What happens when Alice and Bob that are in a relative motion try to communicate? Assume that they use qubits that were described above. Under a general Lorentz transformation Λ that relates Alice's and Bob's frames, the state of this qubit will be transformed due to three distinct effects, which are:

(i) A Wigner rotation due to the Lorentz boost Λ, which occurs even for momentum eigenstates. If Λ is known, then to the extent that the wave-packet spread can be ignored, this is inconsequencial.

(ii) A decoherence due to the entangling of spin and momentum under the Lorentz transformation Λ because the particle is not in a momentum eigenstate. Although reduced or effective density matrices have no general transformation rule, such rules can be established for particular classes of experimental procedures. We can then ask how these effective transformation rules, $\rho' = T(\rho)$, fit into the framework of general state transformations. E.g., for the massive qubit of Sec. 4.2 the effective transformation is given by

$$\rho' = \rho(1 - \frac{\Gamma^2}{4}) + (\sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y)\frac{\Gamma^2}{8}. \tag{79}$$

If Λ is known and it is possible to implement the operators that were mentioned in the Remark 1 above, then this effect is absent. Otherwise, this noise is unavoidable. Still, it is worth to keep in mind that the motion can improve the message fidelity, as in Eq. (78).

(iii) Another kind of decoherence arises due to Bob's lack of knowledge about the transformation relating his reference frame to Alice's frame. Using the techniques of the decoherence-free subspaces, it is possible to eliminate this effect completely. E. g., for massive particles four physical qubits may be used to encode a logical qubit, while for photons $2 \to 1$ encoding is sufficient. In both cases using the block encoding it is possible to reach an asymptotically unit efficiency [36].

Entanglement between the "qubit" and spatial degrees of freedom leads to an interesting complication of the analysis. It is known [1], that the dynamics of a subsystem my be not completely positive, there is a prior entanglement with another system and the dynamics is not factorizable. Since in Eq. (78) and in the discussion following Eq. (54) we have seen that distinguishability *can* be improved, we conclude that these transformations are *not completely positive*. The reason is that the Lorentz transformation acts not only on the "interesting" discrete variables, but also on the primary momentum variables that we elected to ignore and to trace out, and its action on the interesting degrees of freedom depends on the "hidden" primary ones. Of course, the complete state, with all the variables, transforms unitarily and distinguishability is preserved.

This technicality has one important consequence. In quantum information theory quantum channels are described by completely positive maps that act on qubit states. Qubits themselves are realized as discrete degrees of freedom of various particles. If relativistic motion is important, then not only does the vacuum behave as a noisy quantum channel, but the very representation of a channel by a CP map fails.

6. Entanglement and Different Lorentz Observers

In this section I consider only two-particle states. Even in this simple setting there are several possible answers to the question what happens to the entanglement, depending on the details of the question. Since the quantum Lorentz transformation is given by a tensor product $U_1(\Lambda)\otimes U_2(\Lambda)$, the overall entanglement between the states is Lorentz-invariant.

Let us assume that the states can be approximated by momentum eigenstates. Then, the same conclusion applies to the spin-spin (or polarization-polarization) entanglement between the particles, and it is possible to write an appropriate entanglement measures that capture the effects of particle statistics and Lorentz-invariance of the entanglement [23]. However, it does not mean that this invariance will be observed in an experiment, or that the violation of Bell-type inequalities that is observed in an experiment that is performed in Alice's frame will be observed if the same equipment is placed in Bob's frame.

While a field-theoretical analysis shows that violations of Bell-type inequalities are generic, there are conditions that are imposed on the experimental procedures that are used to detect them. Consider the CHSH inequality. For any two spacelike separated regions and any pairs of of operators, a, b, there is a state ρ such that the CHSH inequality is violated, i.e., $\zeta(a, b, \rho) > 1$. With additional technical assumptions the existence of a maximally violating state ρ_m can be proved:

$$\zeta(a, b, \rho_m) = \sqrt{2}, \tag{80}$$

for any spacelike separated regions $\mathcal{O}_L$ and $\mathcal{O}_R$. It follows from convexity arguments that states that maximally violate Bell inequalities are pure. What are then the operators that lead to the maximal violation? It was shown [37] that the operators A_j and B_k that give $\zeta = \sqrt{2}$ satisfy $A_j^2 = \mathbb{1}$ and $A_1A_2 + A_2A_1 = 0$, and likewise for B_k. If we define $A_3 := -i[A_1, A_2]/2$, then these three operators have the same algebra as Pauli spin matrices.

In principle the vacuum state may lead to the maximal violation of Bell-type inequalities. Their observability was discussed in [39].

The operators [38]

$$A_i = 2\left[\frac{m}{p^0}\mathbf{a}_i + \left(1 - \frac{m}{p^0}\right)(\mathbf{a}\cdot\mathbf{n})\mathbf{n}\right]\cdot\mathbf{S} \equiv 2\boldsymbol{\alpha}(\mathbf{a},\mathbf{p})\cdot\mathbf{S}, \tag{81}$$

where $\mathbf{S}$ is the Wigner spin operator and $\mathbf{n} = \mathbf{p}/|\mathbf{p}|$ appear quite naturally as the candidates for the measurement description. The length of the auxiliary vector $\boldsymbol{\alpha}$ is

$$|\boldsymbol{\alpha}| = \frac{\sqrt{(\mathbf{p}\cdot\mathbf{a})^2 + m^2}}{p^0}, \tag{82}$$

so generically $A_i^2 = \boldsymbol{\alpha}^2\mathbb{1} < \mathbb{1}$, and indeed, the degree of violation decreases with the velocity of the observer. Nevertheless, it is always possible to compensate for a Wigner rotation by an appropriate choice of the operators [1].

Realistic situations involve wave packets. For example, a general spin-$\frac{1}{2}$ two-particle state may be written as

$$|\Upsilon_{12}\rangle = \sum_{\sigma_1,\sigma_2}\int d\mu(p_1)d\mu(p_2)g(\sigma_1\sigma_2,\mathbf{p}_1,\mathbf{p}_2)|\mathbf{p}_1,\sigma_1;\mathbf{p}_2,\sigma_2\rangle. \tag{83}$$

For particles with well defined momenta, g sharply peaks at some values $\mathbf{p}_{10}, \mathbf{p}_{20}$. Again, a boost to any Lorentz frame S' will result in a unitary $U(\Lambda)\otimes U(\Lambda)$, acting on each particle separately, thus preserving the entanglement. Nevertheless, since they can change entanglement between different degrees of freedom of a single particle, the spin-spin entanglement is frame-dependent as well. Having investigated the reduced density matrix for $|\Upsilon_{12}\rangle$ and made explicit calculations for the case where g is a Gaussian, as in the Sec. 4.2 above, it is possible to show that if two particles are maximally entangled in a common (approximate) rest frame (Alice's frame), then the concurrence, as seen by a Lorentz-boosted Bob, decreases when $v \to 1$. Of course, the inverse transformation from Bob to Alice will increase the concurrence [40]. Thus, we see that that spin-spin entanglement is not a Lorentz invariant quantity, exactly as spin entropy is not a Lorentz scalar. Relativistic properties of the polarization entanglement are even more interesting [28], since there is no frame where polarization and momentum are unentangled.

7. The Omissions & Perspectives

Because of the lack of space I am only going to mention the various fascinating areas of the interplay between quantum information theory and relativistic physics. Quantum field theory provides us with new situations that should be investigated. For example, it is possible to ask all the usual questions about entanglement, distillability, etc and their invariance[1,41,42,43]. So far we discussed only observers that move with constant velocity. An accelerated observer sees Unruh radiation. It leads to a host of interesting effects if we consider a teleportation between a stationary and accelerated observer [44,45,46]. Dynamical entanglement — the one appearing in the scattering processes or between the decay products also have been investigated [24,47,48].

Going to more exotic settings, I just mention that black hole physics, cosmology, loop quantum gravity and string theory provide extremely interesting scenarios where the questions of information can and should be asked[1,49,50,51,52,53,54]

References

[1] A. Peres and D. R. Terno, *Rev. Mod. Phys.* **76** (2004), 93.
[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, New York, 2000.
[3] M. Keyl, Phys. Rep. **369** (2002), 431.
[4] P. Busch, M. Grabowski, and P. J. Lahti, *Operational Quantum Physics*, Springer, Berlin, 1995.
[5] H.-P. Breuer and F. Petruccione (eds.), *Relativistic Quantum Measurement and Decoherence*, Springer, Berlin, 2000.
[6] H.-P. Breuer (ed.), *Relativistic Quantum Measurement*, Springer, Berlin, 2000.
[7] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, *Phys. Rev. A* **59** (1999), 1070.
[8] D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskill, *Phys. Rev. A* **64** (2001), 052309.
[9] T. Eggeling, Schlingemann, and R. F. Werner, *Europhys. Lett.* **57** (2002), 782.
[10] B. Groisman and B. Reznik, *Phys. Rev. A* **66**, (2002), 022110.
[11] L. Vaidman, *Phys. Rev. Lett.* **90**, (2003) 010402.
[12] M. Florig and S. J. Summers, *J. Math. Phys.* **38** (1997), 1318.
[13] A. Kent, *Phys. Rev. Lett.* **83** (1999), 1447.
[14] A. Kent, *Phys. Rev. Lett.* **90** (2003), 237901.
[15] O. Keller, *Phys. Rep.* **411** (2005), 1.
[16] M. Toller, *Phys. Rev. A*, **59** (1999), 960.
[17] S. Mazzucchi, *J. Math. Phys.* **42** (2001), 2477.
[18] W.-K. Tung, *Group Theory in Physics*, World Scientific, Singapore, 1985.
[19] S. Weinberg, *The Quantum Theory of Fields* Vol. 1, Cambridge University, Cambridge, 1995.
[20] A. O. Barut and R. Raczka, *Theory of Group Representations and Applications*, World Scientific, Singapore, 1987.
[21] N.N. Bogolubov, A. A. Logunov, A. I. Oksak, and I. T. Todorov, *General Principles of Quantum Field Theory*, Kluwer, Dordrecht, 1990.
[22] F. R. Halpern, *Special Relativity and Quantum Mechanics*, Prentice-Hall, Englewood Cliffs, 1968.
[23] C. Soo and C. C. Y. Lin, *Int. J. Quant. Info.* **2** (2004), 183.
[24] N. L. Harshman, *Phys. Rev. A* **71** (2005), 022312.
[25] D. R. Terno, *Phys. Rev. A* **67** (2003), 014102.
[26] P. M. Alsing and G. J. Milburn, *Quant. Info. Comp.* **2**, (2002), 487.
[27] N. H. Lindner, A. Peres, and D. R. Terno, *J. Phys. A* **36** (2003), L449.
[28] A. J. Bergou, R. M. Gingrich, and C. Adami, *Phys. Rev. A* **68** (2003), 042102.
[29] A. Peres and D. R. Terno, *J. Mod. Opt.* **50** (2003), 1165.
[30] A. Aiello and J. P. Woerdman, *Phys. Rev. A* **70** (2004), 023808.
[31] N. H. Lindner and D. R. Terno, *J. Mod. Opt.* **52** (2005), 1177.
[32] M. Czachor and M. Wilczewski, *Phys. Rev. A* **68** (2003), 010302(R).
[33] M. Czachor, *Phys. Rev. Lett.* **94** (2005), 078901.
[34] P. Kok, T. C. Ralph, and G. J. Milburn, *Quant. Info. Comp.* **5** (2005), 239.
[35] P. Kok and S. L. Braunstein, *Int. J. Quant. Info.* **4** (2006), in press.
[36] S. D. Bartlett and D. R. Terno, *Phys. Rev. A* **71** (2005), 012302.
[37] S. J. Summers and R. Werner, *J. Math. Phys.* **28** (1987), 2440.
[38] M. Czachor, *Phys. Rev. A* **55** (1997), 72.
[39] B. Reznik, A. Retzker and J. Silman, *Phys. Rev. A* **71** (2005), 042104.
[40] R. M. Gingrich and C. Adami, *Phys. Rev. Lett.* **89** (2002), 270402.
[41] R. Verch and R. F. Werner, *Rev. Math. Phys.* **17** (2005), 545.
[42] D. R. Terno, *Phys. Rev. Lett.* **93** (2004) 051303.
[43] M. B. Plenio, J. Eisert, J. Dreissig, and M. Cramer *Phys. Rev. Lett.* **94** (2005), 060503.

[44] P. M. Alsing and G. J. Milburn, *Phys. Rev. Lett.* **91** (2003), 180404.
[45] I. Fuentes-Schuller and R. B. Mann, *Phys. Rev. Lett.* **95** (2005), 120404.
[46] R. Schützhold and W. G. Unruh, e-print quant-ph/0506028.
[47] E. B. Manoukian and N. Yongram, *Eur. J. Phys. D* **31** (2004), 137.
[48] L. Lamata, J. Leon, and E. Solano, e-print quant-ph/0509021.
[49] D. Campo and R. Parentani, e-print astro-ph/0505376.
[50] O. Dreyer, F. Markopoulou, and L. Smolin, e-print hep-th/0409056.
[51] E. R. Livine and D. R. Terno, *Phys. Rev. A* **72** (2005), 022307.
[52] E. R. Livine and D. R. Terno, e-print gr-qc/0508085.
[53] D. R. Terno, e-print gr-qc/0505068.
[54] E. B. Manoukian and N. Yongram, *Mod. Phys. Lett. A* **20** (2005), 623.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Generalized Bell Inequalities and the Entanglement of Pure States

Kwek Leong-Chuan [a,b,1], Chunfeng Wu [b], Jingling Chen [c], Dagomir Kaszlikowski [b] and C.H. Oh [b]

[a] *National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 639798*

[b] *Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542*

[c] *Theoretical Physics Division, Chern Institute of Mathematics, Nankai University, Tianjin 300071, P. R. China*

Abstract. It is well known that all pure entangled bipartite states violate Bell inequalities. When one extends Bell inequalities to multipartite two dimensional states using either the Mermin inequalities or Werner-Wolf-Żukowski-Brukner inequalities, one finds that it is no longer true that all entangled pure states violate these inequalities. We show that it is possible to formulate a set of Bell inequalities for 3-qubit system in such a way that all entangled tripartite states violate it.

1. Introduction

Bell inequalities were originally devised by John Bell in 1964 to demonstrate the possible violation of local realistic theories for quantum mechanical systems. It was subsequently shown by Gisin that any pure entangled state of two particles violates any standard Bell inequality [1][2]. By a standard Bell experiment we mean the one in which each local observer is given a choice between two dichotomic observables [3][4]. For mixed states, the situation is more complex and there exist entangled mixed bipartite states that do not violate Bell inequalities.

In quantum information theory, an entangled state is an important and useful resource for both computation and communication. Since Bell inequalities can detect pure entangled states, there has been renewed interest in Bell inequalities and their generalizations. In particular, Mermin-Ardehali-Belinskii-Klyshko (MABK) devised a set of inequalities for $N > 2$ qubits [5]. Subsequently, two independent groups, Werner-Wolf and Żukowski-Brukner [4] derived a general set of Bell inequalities, also known as WWZB inequalities, for correlation functions for N qubits. The WWZB inequalities include MABK inequalities as special cases. Generalizations to higher dimensional systems have also been investigated, namely to two $d(d > 2)$-dimensional systems [6,7,8,9] as well as to $N > 2$ $d(d > 2)$-dimensional systems [10].

[1]Correspondence to: L.C. Kwek, 1, Nanyang Walk, 7-03-86, Singapore 637616. Tel.: +65 6790-3823; Fax: +65 222 333 0000; E-mail: lckwek@nie.edu.sg.

Unfortunately, for WWZB inequalities, it turns out that there exists a family of pure entangled $N > 2$ qubit states that do not violate any Bell inequality for N-particle correlations for the case of a standard Bell experiment on N qubits [11]. This family is the generalized GHZ states given by

$$|\psi\rangle_{GHZ} = \cos\xi|0\cdots 0\rangle + \sin\xi|1\cdots 1\rangle \tag{1}$$

with $0 \leq \xi \leq \pi/4$. Note that the GHZ states [12] corresponds to $\xi = \pi/4$. Moreover, Scarani and Gisin noticed that for $\sin 2\xi \leq 1/\sqrt{2^{N-1}}$ the states (1) do not violate the Mermin-Ardehali-Belinskii-Klyshko (MABK) inequalities. Indeed, Ref. [3] shows that (a) For $N = \text{even}$, although the generalized GHZ state (1) does not violate MABK inequalities, it violates the ZB inequality and (b) For $\sin 2\xi \leq 1/\sqrt{2^{N-1}}$ and $N = \text{odd}$, the correlations between measurements on qubits in the generalized GHZ state (1) satisfy all standard Bell inequalities.

2. Three-Qubit System

For a three-qubit system, a natural question in this context is: Can Gisin's theorem be generalized to 3-qubit pure entangled states? Can one find a Bell inequality that violates $|\psi\rangle_{GHZ}$ for all pure entangled tripartite states?

Before proceeding further, we need to consider the complete classification of 3-qubit entanglement. Entanglement for N-qubit systems using a quadratic Bell inequality based on MABK inequalities has been presented in Ref. [13]. For $N = 3$, there are three different types of 3-qubit states: (i) totally separable states denoted as (1_3)={ mixtures of states of form $\rho_A \otimes \rho_B \otimes \rho_C$}; (ii) 2-entangled states which are denoted as $(2,1)$={mixtures of states of form $\rho_A \otimes \rho_{BC}, \rho_{AC} \otimes \rho_B, \rho_{AB} \otimes \rho_C$}; (iii) fully entangled states which are denoted as $(3) = \{\rho_{ABC}\}$ including the GHZ state. Thus we see that there are two different types of entangled pure 3-qubit states, namely, the 2-entangled states and the fully entangled states.

To construct the Bell inequalities for 3-qubit case, let us consider the following Bell-type scenario: three space-separated observers, denoted by A, B and C (or Alice, Bob and Charlie), who can measure two different local observables with two outcomes, labeled by 0 and 1. We denote X_i as the observable measured by party X and x_i as the outcome, where $X = A, B, C$ $(x = a, b, c)$. If the observers decide to measure A_1, B_1 and C_2, the result is $(0, 1, 1)$ with probability $P(a_1 = 0, b_1 = 1, c_2 = 1)$. In Ref. [14], we introduced the Bell inequality with all possible probabilities as:

$$\begin{aligned}
&P(a_1 + b_1 + c_1 = 1) + 2P(a_2 + b_2 + c_2 = 1) + P(a_1 + b_2 + c_2 = 2) \\
&+P(a_2 + b_1 + c_2 = 2) + P(a_2 + b_2 + c_1 = 2) - P(a_1 + b_1 + c_2 = 0) \\
&-P(a_1 + b_2 + c_1 = 0) - P(a_2 + b_1 + c_1 = 0) - P(a_1 + b_1 + c_2 = 3) \\
&-P(a_1 + b_2 + c_1 = 3) - P(a_2 + b_1 + c_1 = 3) \leq 3.
\end{aligned} \tag{2}$$

This inequality is symmetric under the permutations of three observers Alice, Bob and Charlie. Pure states of three qubits constitute a five-parameter family, with equivalence up to local unitary transformations. This family has the representation [10]

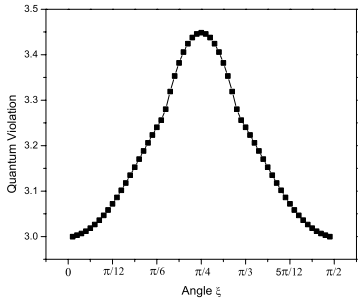


Figure 1. Numerical results for the generalized GHZ states $|\psi\rangle_{GHZ} = \cos\xi|000\rangle + \sin\xi|111\rangle$, which violate Bell inequality for probabilities (2) except $\xi = 0$ and $\pi/2$. For the GHZ state with $\xi = \pi/4$, the Bell quantity reaches its maximum value $\frac{3}{8}(4 + 3\sqrt{3})$.

$$|\psi\rangle = \sqrt{\mu_0}|000\rangle + \sqrt{\mu_1}e^{i\phi}|100\rangle + \sqrt{\mu_2}|101\rangle + \sqrt{\mu_3}|110\rangle + \sqrt{\mu_4}|111\rangle \tag{3}$$

with $\mu_i \geq 0$, $\sum_i \mu_i = 1$ and $0 \leq \phi \leq \pi$. Numerical results show that this Bell inequality for probabilities is violated by all pure entangled states of three-qubit system even though it is hard to provide an analytic proof.

3. Numerical Results

In Fig.1, we show the numerical results for the generalized GHZ states $|\psi\rangle_{GHZ} = \cos\xi|000\rangle + \sin\xi|111\rangle$, which violate the above symmetric Bell inequality for probabilities except $\xi = 0$ and $\pi/2$. For the measuring angles $\theta_{a_1} = \theta_{a_2} = \theta_{b_1} = \theta_{b_2} = \theta_{c_1} = \theta_{c_2} = \pi/2, \phi_{a_1} = -5\pi/12, \phi_{a_2} = \pi/4, \phi_{b_1} = -5\pi/12, \phi_{b_2} = \pi/4, \phi_{c_1} = -\pi/3, \phi_{c_2} = \pi/3$, all the probability terms with positive signs in Bell inequality (2) are equal to $\frac{3}{16}(2 + \sqrt{3})$, while the terms with negative signs are equal to $\frac{1}{8}$, so the quantum violation of Bell quantity for the GHZ state (where $\xi = \pi/4$) is obtained as $6\times\frac{3}{16}(2+\sqrt{3})-6\times\frac{1}{8} = \frac{3}{8}(4+3\sqrt{3}) > 3$. In Fig.2, we show the numerical results for the family of generalized W states $|\psi\rangle_W = \sin\beta\cos\xi|100\rangle + \sin\beta\sin\xi|010\rangle + \cos\beta|001\rangle$ with the cases $\beta = \pi/12, \pi/6, \pi/4, \pi/3, 5\pi/12$ and $\pi/2$, which show the quantum violation of $|\psi\rangle_W$ except the product cases with $\beta = \pi/2, \xi = 0$ and $\pi/2$. For the standard W state $|\psi\rangle_W = (|100\rangle + |010\rangle + |001\rangle)/\sqrt{3}$, the quantum violation is 3.55153. We now proceed to present the second theorem.

For pure 2-entangled states of three-qubit system, we need to consider the following: $|\psi_{AB}\rangle \otimes |\psi_C\rangle$, $|\psi_{AC}\rangle \otimes |\psi_B\rangle$ and $|\psi_{BC}\rangle \otimes |\psi_A\rangle$. It is however sufficient to consider only one of them, say $|\psi_{AB}\rangle \otimes |\psi_C\rangle$, since Bell inequality (2) is symmetric under the permutations of A, B and C. Moreover, one can always have $|\psi_{AB}\rangle \otimes |\psi_C\rangle = (\cos\xi|00\rangle_{AB} + \sin\xi|11\rangle_{AB}) \otimes |0\rangle_C$ due to local unitary transformations. For the measuring angles $\theta_{a_1} = \theta_{a_2} = \theta, \phi_{a_1} = 2\pi/3, \phi_{a_2} = -\pi/3, \theta_{b_1} = \theta_{c_1} = 0, \phi_{b_1} = \phi_{c_1} = 0, \theta_{b_2} = \pi/2, \theta_{c_2} = \pi, \phi_{b_2} = \pi/3, \phi_{c_2} = 0$, we obtain from the left-hand side of Bell inequality (2) that

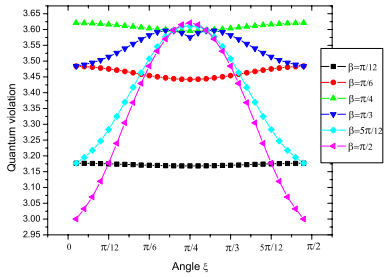


Figure 2. Numerical results for the family of generalized W states $|\psi\rangle_W = \sin\beta\cos\xi|100\rangle + \sin\beta\sin\xi|010\rangle + \cos\beta|001\rangle$ with the cases $\beta = \pi/12, \pi/6, \pi/4, \pi/3, 5\pi/12$ and $\pi/2$.

$$\begin{aligned}\mathcal{B} &= \frac{3}{2}(1 - \cos\theta + \sin(2\xi)\sin\theta) \\ &\geq \frac{3}{2}(1 + \sqrt{1 + \sin^2(2\xi)}),\end{aligned} \tag{4}$$

where equality occurs at $\theta = -\tan^{-1}[\sin(2\xi)]$. Obviously the Bell inequality is violated for any $\xi \neq 0$ or $\pi/2$. This ends the proof. Indeed, the quantum violation of the state $|\psi_{AB}\rangle\otimes|\psi_C\rangle$ corresponds to the curve with $\beta = \pi/2$ as shown in Fig.2, because $|\psi_{AB}\rangle\otimes|\psi_C\rangle$ is equivalent to $|\psi\rangle_W$ for $\beta = \pi/2$ up to a local unitary transformation.

There is a simpler and more intuitive way to show that 2-entangled states violate the three-qubit Bell inequality: the symmetric Bell inequality (2) can be reduced to a CHSH-like inequality for two qubits and then from Gisin's theorem for two qubits one can easily deduce the result. Indeed, by taking $c_1 = 0, c_2 = 1$, we have from Eq. (2) that

$$\begin{aligned}&P(a_1 + b_1 = 1) + 2P(a_2 + b_2 = 0) \\ &+P(a_1 + b_2 = 1) + P(a_2 + b_1 = 1) + P(a_2 + b_2 = 2) \\ &-P(a_1 + b_1 = -1) - P(a_1 + b_2 = 0) - P(a_2 + b_1 = 0) \\ &-P(a_1 + b_1 = 2) - P(a_1 + b_2 = 3) - P(a_2 + b_1 = 3) \leq 3\end{aligned} \tag{5}$$

Since $a_1, a_2, b_1, b_2 = 0, 1$, the probabilities $P(a_1 + b_1 = -1)$, $P(a_1 + b_2 = 3)$ and $P(a_2 + b_1 = 3)$ will be equal to zero. By using $P(a_2 + b_2 = 0) + P(a_2 + b_2 = 2) = 1 - P(a_2 + b_2 = 1)$, we arrive at the following Bell inequality for two-qubits:

$$\begin{aligned}&P(a_1 + b_1 = 1) + P(a_1 + b_2 = 1) + P(a_2 + b_1 = 1) \\ &+P(a_2 + b_2 = 0) - P(a_1 + b_1 = 2) - P(a_1 + b_2 = 0) \\ &-P(a_2 + b_1 = 0) - P(a_2 + b_2 = 1) \leq 2.\end{aligned} \tag{6}$$

This Bell inequality is symmetric under the permutations of Alice and Bob, it is an alternative form for CHSH inequality of two qubits. For the two-qubit state $|\psi\rangle = \cos\xi|00\rangle + \sin\xi|11\rangle$, one can have the quantum probability

$$
\begin{aligned}
&P^{QM}(a_i = m, b_j = n) \\
&= \frac{1}{4}\cos^2\xi[1 + (-1)^m \cos\theta_{a_i}][1 + (-1)^n \cos\theta_{b_j}] \\
&+ \frac{1}{4}\sin^2\xi[1 - (-1)^m \cos\theta_{a_i}][1 - (-1)^n \cos\theta_{b_j}] \\
&+ \frac{1}{4}\sin(2\xi)(-1)^{m+n} \sin\theta_{a_i} \sin\theta_{b_j} \cos(\phi_{a_i} + \phi_{b_j})
\end{aligned} \tag{7}
$$

For the measuring angles $\theta_{a_1} = \theta_{a_2} = \theta, \phi_{a_1} = \pi - \phi, \phi_{a_2} = -\phi, \theta_{b_1} = 0, \phi_{b_1} = 0, \theta_{b_2} = \pi/2, \phi_{b_2} = \phi$, the left-hand side of Bell inequality (6) becomes $\mathcal{B} = \frac{1}{2} + \frac{3}{2}(-\cos\theta + \sin(2\xi)\sin\theta) \geq \frac{1}{2}(1 + 3\sqrt{1 + \sin^2(2\xi)})$, where equality occurs at $\theta = -\tan^{-1}[\sin(2\xi)]$. Obviously the Bell inequality (6) is violated for any $\xi \neq 0$ or $\frac{\pi}{2}$, just as the CHSH inequality is violated by the 2-qubit state $|\psi\rangle = \cos\xi|00\rangle + \sin\xi|11\rangle$. For the Werner state $\rho_W = V|\psi\rangle\langle\psi| + (1 - V)\rho_{\rm noise}$, where $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ is the maximally entangled state. The maximal value of V for which local realism is still possible by this Bell inequality is $V_{max} = 1/\sqrt{2}$, just the same as the case for a CHSH inequality. Actually, if one denotes the left-hand side of Bell inequality (6) by $\mathcal{B}$ and redefines a new Bell quantity $\mathcal{B}' = \frac{4}{3}(\mathcal{B} - \frac{1}{2})$, you still have the Bell inequality $\mathcal{B}' \leq 2$. For quantum mechanics, $\mathcal{B}'_{max} = 2\sqrt{1 + \sin^2(2\xi)}$, which reaches $2\sqrt{2}$ and then $\mathcal{B}'$ recovers the usual CHSH inequality.

4. Summary

In summary, (i) since all pure entangled states (including pure 2-entangled states) of a three-qubit system violate our Bell inequality (2), we have Gisin's theorem for a 3-qubit system; (ii) the Bell inequality (2) can be reduced to an alternative form of the CHSH inequality (in terms of probabilities), thus it can be viewed as a good candidate for a"natural" generalization of the usual CHSH inequality. (iii) MABK inequalities and WWZB inequalities are binary correlation Bell inequalities. However, one may notice that the Bell inequality (2) is a ternary Bell inequalities, i.e., where the inequalities are "modulo 3". Note that the three-qutrit inequality [10] can be connected to Bell inequality (2), which is for three qubits, if one restricts the initial three possible outcomes of each measurement to only two possible outcomes.

5. Acknowledgment

L.C. Kwek would like to thank A.Ekert, D. Angelakis and S. Kulik for their hospitality during the NATO-ASI conference in Chania, Crete.

References

[1] N. Gisin, Phys. Lett. A **154**, 201 (1991); N. Gisin and A. Peres, Phys. Lett. A **162**, 15-17 (1992).

[2] S. Popescu and D. Rohrlich, Phys. Lett. A **166**, 293 (1992).
[3] M. Żukowski, Č. Brukner, W. Laskowski, and M. Wiesniak, Phys. Rev. Lett. **88**, 210402 (2002).
[4] M. Żukowski and Č. Brukner, Phys. Rev. Lett. **88**, 210401 (2002).
[5] N. D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990); M. Ardehali, Phys. Rev. A **46**, 5375 (1992); A. V. Belinskii and D. N. Klyshko, Phys. Usp. **36**, 653 (1993).
[6] D. Kaszlikowski, L.C. Kwek, J.L. Chen, M. Żukowski and C.H. Oh, Phys. Rev. A, **65** (2002) 032118.
[7] D. Collins, N. Gisin, N. Linden, S. Massar and S. Popescu, quant-ph//0106024.
[8] D. Kaszlikowski, P. Gnaciński, M. Żukowski, W. Miklaszewski and A. Zeilinger, *Phys. Rev. Lett.* **85** (2000) 4418.
[9] T. Durt, D. Kaszlikowski and M. Żukowski, Phys. Rev. A **64** (2001) 024101.
[10] A. Acin, J.L. Chen, N. Gisin, D. Kaszlikowski, L.C. Kwek, C.H. Oh and M. Zukowski, *Phys. Rev. Lett.*, **92** (2004) 250404
[11] V. Scarani and N. Gisin, J. Phys. A **34** 6043 (2001).
[12] D. M. Greenberger, M. Horne, A. Shimony, and A. Zeilinger, Am. J. Phys. **58**, 1131 (1990).
[13] X. Yu, Z.B. Chen, J.W. Pan, and Y.D. Zhang, e-print quant-ph/0112102.
[14] Jingling Chen, Chunfeng Wu, L.C. Kwek and C.H. Oh, *Phys. Rev. Lett.* **93 (2004) 140407.**

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Thermal Entanglement in Infinite Dimensional Systems

A. Ferreira [a,1], A. Guerreiro [a] and V. Vedral [b]

[a] *Dep. Física Fac. Ciências Univ. Porto, Campo Alegre 687, 4169-007, Porto, Portugal*
[b] *The School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT, UK*

Abstract. We demonstrate the persistence of entanglement at any temperature in a system composed by a cavity field mode interacting with a movable mirror, via radiation pressure coupling. We introduce a new method of analyzing entanglement in mixed states of infinite dimensional systems and identify an important symmetry between a special class of 2×2 subspaces of the density matrix.

Keywords. Entanglement, Thermal entanglement

1. Introduction

Since the development of quantum theory, much interest exists in understanding the transition between the quantum and the classical behavior of physical systems. Different schemes to create and probe macroscopical quantum superpositions have already been proposed for different physical systems [1][2].

In this paper we proceed in a different direction in order to understand how quantum mechanics can accomplish entanglement in the macroscopic domain. It is a general believe that temperature destroys the entanglement since the system becomes more mixed at higher temperatures. In [3] it has been shown that every finite dimensional bipartite system in equilibrium with a thermal bath has a critical temperature above which entanglement is no longer present. Also it is well known that a mixture of a maximally mixed state and a Bell state of two qubits, $p|\psi_-\rangle\langle\psi_-| + (1-p)I/4$, the so called Werner state [4], will only be entangled for $0 < p < 2/9$. This means that the increase of the mixedness parameter, p, like the increase of the temperature, will make the state separable at some point.

We will consider a well studied physical system [1] [5] composed by a movable mirror coupled to a cavity electromagnetic field. A new method of analyzing entanglement in an infinite dimensional bipartite system will be introduced and it will be shown that, in deep contrast with a finite dimensional system, infinite dimensional systems are able to produce entanglement at arbitrarily high temperatures through radiation pressure interaction.

[1]Correspondence to: Aires F. Ferreira, Departamento de Física da Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal; Tel.: +359 22 608 2611; Fax: +359 608 2628; E-mail:aires.ferreira@univie.ac.at

2. The System Evolution

We will consider a perfect cavity with a movable mirror in one end, modeled as a mechanical harmonic oscillator, whose quivering motion is quantized. A driving laser tuned to resonance with a cavity mode prepares a coherent state of the electromagnetic field $|\alpha\rangle$.

When the mirror is considered the initial state of the system is described by a thermal state rather than a pure one. This can be done in the standard way by writing the coherent state decomposition,

$$\rho(0) = \frac{1}{\mathcal{Z}} \int \frac{d^2z}{\pi} e^{-|z|^2/\bar{n}} |\alpha\rangle\langle\alpha| \otimes |z\rangle\langle z|, \tag{1}$$

where $\bar{n} = 1/(e^{\hbar w_m/K_B T} - 1)$ is the mean number of excitations at some temperature T, $\mathcal{Z}$ is the partition function for the mirror with resonance frequency w_m, α is the input coherent light state amplitude and z represents all the possible coherent states of the mirror.

The Hamiltonian of this system can be expressed in a simple way by considering only the free Hamiltonian plus the interaction term due to radiation pressure, i.e. neglecting the scattering from the driven cavity and retardation effects due to the mirror oscillation. This corresponds to the adiabatic limit which is easily achieved for optical frequencies of the cavity field. Hence, neglecting photon losses in the cavity, the Hamiltonian reads [6]:

$$H = \hbar w_0 a^\dagger a + \hbar w_m b^\dagger b - \hbar g a^\dagger a (b + b^\dagger), \tag{2}$$

where a is the annihilation operator of the cavity mode, b is the annihilation operator of the mirror, w_0 is the frequency of the cavity mode and $g = w_0\sqrt{\hbar}/(L\sqrt{m w_m})$ is the coupling constant.

In the interaction picture, the evolution operator associated with the Hamiltonian (2) has a closed formula and it was derived in [1], using the Campbell-Baker-Hausdorff formula for the Lie algebra, and in [9], using operator algebra methods: $U(t) = e^{ik^2(a^\dagger a)^2\Lambda(t)} D_m(\eta(t) k a^\dagger a)$, where $\eta(t) = 1 - e^{-iw_m t}$, $\Lambda(t) = w_m t - sin(w_m t)$, $k = g/w_m$ and D_m is the displacement operator of the mirror, $D_m(\gamma)|0\rangle = |\gamma\rangle$ with $\gamma \in \mathcal{C}$. The evolution of the system is periodic and we only need to investigate entanglement in the time interval $[0, 2\pi/w_m]$.

When the general case is considered, the density matrix, $\rho(t)$ expressed in the Fock states of both cavity field and mirror reads:

$$\rho(t) = \sum_{\mu,\nu,n,m} \rho_{\mu\nu nm}(t) |n\rangle\langle m| \otimes |\mu\rangle\langle\nu|, \tag{3}$$

where the Latin indexes refer to the radiation and the Greek indexes refer to the mirror. By applying the evolution operator $U(t)\rho(0)U(t)^\dagger$ to the initial state (1), the elements of the density matrix are calculated to be:

$$\rho_{\mu\nu nm}(t) = \Phi_{nm\mu\nu}(t) \int \frac{d^2z}{\pi} e^{-|z|^2/\overline{n}} F_n(t)^\mu F_m(t)^\nu e^{-\frac{|F_n(t)|^2 + |F_m(t)|^2}{2}}, \tag{4}$$

where $\Phi_{nm\mu\nu}(t) = \alpha^n \alpha^{*m} e^{i\Lambda(t)(n^2 - m^2) - |\alpha|^2} / (\mathcal{Z}\sqrt{n!m!\mu!\nu!})$ and $F_n(t) = z + kn\eta(t)$.

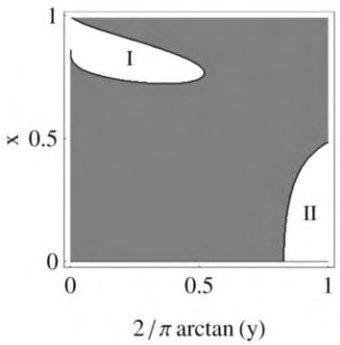


Figure 1. Υ as a function of $(2/\pi)\arctan(y)$ and x for the subspace with $0, 1$ excitations. Entanglement (gray) exists for all parameter space except in regions I and II (white).

3. Entanglement Characterization

Quantifying entanglement in mixed states is a non trivial problem, except for bipartite two level systems, where the Peres criterium is both sufficient [7] and necessary [8] to guarantee entanglement. In this paper a new approach inspired by [9] is followed. First, we project the original density matrix (2) into a 2×2 subspace, which corresponds to a local action thus not increasing the amount of entanglement $E(\rho)$ in the overall system [10], i.e. $E(\rho) \geq E(P_{2\times 2}\rho P_{2\times 2})$. If entanglement is found in the 2×2 dimensional subspace then the density matrix $\rho(t)$ is non separable.

According to the Peres criterium, the existence of entanglement in each 2×2 subspace can be identified by verifying whether the partial transposed density matrix ρ^{T_P} has a negative eigenvalue. For our purposes, it is then sufficient to check if the determinant of ρ^{T_P} is negative.

For the subspace with $m, n = 0, 1$ excitations of the field and $\mu, \nu = 0, 1$ excitations of the mirror, the determinant of ρ^{T_P} is $\Upsilon_{[0,1]}(k) = -G(y,x)H(y,x)$, where $G(y,x) = x^4|\alpha|^4 e^{y^2(x-2)-4|\alpha|^2}/(16\overline{n}^4)$ is always positive, $y \equiv y(t,k) = \sqrt{2}k\sqrt{1-\cos(w_m t)}$ and $x = \overline{n}/(1+\overline{n})$. The sign of Υ is determined by $H(y,x)$ and it is plotted as a function of $(2/\pi)\arctan(y)$ and x in Figure 1, where it is clear that the entanglement occurs for all parameters except for low values of y at high temperatures (region I) and for high values of y and low temperature (region II). In region II the lack of entanglement is not important since, during the evolution of the system, the value of $y(k,t)$ ranges between 0 and $y_{max} = 2k$ and, for sufficiently short times, the system is entangled independently of how large k is. In region I the proof of existence of entanglement is not as evident because, in principle, the system could always be separable for small values of k and high temperatures. However, by noticing that in (4) that k is always multiplied either by n or by m, it is straightforward to verify that $\Upsilon_{[0,s]}$, for the subspace with $m, n = 0, s$ excitations of the field, is proportional to $\Upsilon_{[0,1]}$, for the subspace with $m, n = 0, 1$ excitations of the field, if in the latter the coupling $k_s = k/s$ is chosen, i.e.

$$|\alpha|^{-2s} s! \Upsilon_{[0,s]}(k_s) = |\alpha|^{-2} \Upsilon_{[0,1]}(k). \tag{5}$$

For high temperatures, if we choose a large k capable of producing entanglement in the subspace with $m, n = 0, 1$, then there must be entanglement in the subspace with $m, n = 0, s$ for the coupling constant $k_s = k/s$, even though k_s might not lead to entanglement

in the subspace $m, n = 0, 1$. Therefore, for systems where $y(k, t)$ varies in region II there is always a s such that entanglement occurs in the subspace with $m, n = 0, s$.

4. Conclusion and Final Remarks

In this paper we have shown that the entanglement appears naturally in the dynamics of the system for any value of coupling between the mirror and the cavity field and any finite temperature. Although entanglement occurs at any finite temperature this does not imply the existence of entanglement in the limit of infinite temperature. As expected, in this limit Υ becomes non-negative meaning that entanglement vanishes.

Moreover, we have presented and used a method of identifying entanglement in a physical system with a infinite number of dimensions in a thermal state. This method consists in projecting the density matrix in a 2×2 subspace and verifying entanglement in each subspace. We have identified that the expression of Υ for all the subspaces with $0, 1$ excitations of mirror and with $0, s$ excitations of the cavity field are all formally equivalent. This symmetry implies that the critical temperature (i.e. the temperature above which entanglement is destroyed) for each subspace can be arbitrarily large by allowing higher values of s. Though each 2×2 subspace has a critical temperature , the complete system does not.

The existence of symmetries and patterns in physical expressions is not usually innocent, but a consequence of some fundamental properties of the physical systems. Then, it would not be surprising if similar symmetries appear in other types of systems, thus allowing the same approach of identifying the lack of a critical temperature for the entanglement to be applied. This work suggests that infinite dimensional systems, coupled by some non trivial interaction, can be entangled at any temperature.

Acknowledgements

We gratefully acknowledge helpful discussions with M. Aspelmeyer, Č. Brukner, S. Gigan, J. Kofler, E. Lage, P. Vieira and M. Wieśniak. A.F. is supported by FCT (Portugal) through grant PRAXIS no. SFRH/BD/18092/04. V.V. acknowledges funding from EPSRC and the European Union.

References

[1] S. Mancini *et al.*, Phys. Rev. A **55**, 3042 (1997); S. Bose *et al.*, Phys. Rev. A **56**, 4175 (1997).
[2] W. Marshall *et al*, Phys. Rev. Lett. **91**, 130401 (2003).
[3] B. V. Fine, F. Mintert and A. Buchleitner. Phys. Rev. B **71**, 153105 (2005).
[4] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[5] A. Ferreira, A. Guerreiro and V. Vedral, pre-print: quant-ph/0504186 (2005).
[6] C. K. Law, Phys. Rev. A **51**, 2537 (1995).
[7] A. Peres, Phys. Rev. Letters **77**, 1413 (1996).
[8] M. Horodecki *et al.*, Phys. Lett. A, **223**, 8 (1996).
[9] S. Bose *et al.*, Phys. Rev. Lett. **87**, 050401 (2001).
[10] V. Vedral *et al.* Phys. Rev. Lett. **78**, 2275 (1997);

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Improved Algorithm for Quantum Separability and Entanglement Detection[1]

Lawrence M. Ioannou [a], Benjamin C. Travaglione [a], Donny Cheung [b]
Artur K. Ekert [a]

[a] *DAMTP, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, UK*

[b] *IQC, University of Waterloo, University Avenue, Waterloo, ON, N2L 3G1, Canada*

Abstract. Determining whether a quantum state is separable or entangled is a problem of fundamental importance in quantum information science. It has recently been shown that this problem is NP-hard. There is a highly inefficient 'basic algorithm' for solving the quantum separability problem which follows from the definition of a separable state. By exploiting specific properties of the set of separable states, we introduce a new classical algorithm that solves the problem significantly faster than the 'basic algorithm', allowing a feasible separability test where none previously existed e.g. in 3-by-3-dimensional systems. Our algorithm also provides a novel tool in the experimental detection of entanglement.

Keywords. Quantum separability problem, entanglement detection, complexity, cutting-plane algorithm, interior-point algorithm, analytic center

We begin by introducing some notation and precisely defining the quantum separability problem. In what follows, we are considering a bipartite quantum system of dimension $M \times N$. Let $\mathbf{H}_{M,N}$ denote the vector space of all Hermitian operators acting on $\mathbf{C}^M \otimes \mathbf{C}^N$. Noting that $\mathbf{H}_{M,N}$ is isomorphic to $\mathbf{R}^{M^2N^2}$, it is endowed with the Euclidean inner-product $\langle X, Y\rangle \equiv \mathrm{tr}(XY)$, which induces the corresponding norm $||X|| \equiv \sqrt{\mathrm{tr}(X^2)}$ and distance measure $||X - Y||$. Let $\mathcal{D}_{M,N} \subset \mathbf{H}_{M,N}$ denote the set of all density operators. The set of bipartite separable quantum states $\mathcal{S}_{M,N} \subset \mathcal{D}_{M,N}$ is defined as the convex hull of the separable pure states $\{|\alpha\rangle\langle\alpha| \otimes |\beta\rangle\langle\beta|\}$, where $|\alpha\rangle$ ($|\beta\rangle$) is a normalized vector in $\mathbf{C}^M$ ($\mathbf{C}^N$). An arbitrary density matrix in $\mathcal{D}_{M,N}$ is parameterized by $n - 1$ real variables, where

$$n \equiv M^2N^2, \tag{1}$$

and an arbitrary separable pure state is parameterized by

$$k \equiv 2(M + N) - 4 \tag{2}$$

[1] Excerpt from [17] by permission of Physical Review.

real variables. Of course, in defining the separability problem, we cannot allow infinite precision, so we need to introduce a precision parameter $\delta > 0$. We are now ready to define the *(quantum) separability problem* as follows:

Quantum Separability Problem. *Given a density matrix $\rho \in \mathcal{D}_{M,N}$ and a precision δ, assert either*

SEPARABLE: *there exists a separable state σ such that $||\rho - \sigma|| < \delta$; or*
ENTANGLED: *there exists an entangled state τ such that $||\rho - \tau|| < \delta$.*

The separability problem has been shown to be NP-hard [1], thus any devised test for separability is likely to require a number of computing resources that scales exponentially with M and N. There exist efficient "one-sided" tests for separability, where the output of some polynomial-time computable function of the matrix for ρ can indicate that ρ is certainly entangled [2,3,4,5,6] or certainly separable [7,8,9], but not both.

The experimental flavor of the separability problem can be defined as follows: Given many physical copies of a completely unknown quantum state ρ, determine whether ρ is separable. One may solve this problem by performing full state tomography in order to construct the density matrix for ρ to some precision δ, and then solve the theoretical separability problem. If there is some partial knowledge of ρ, then there are more options, such as testing for a violation of a specific Bell inequality [10,11] or invoking entanglement witnesses [12,13]. In the case where $MN \leq 6$, the positive partial transpose (PPT) test [2,14] can be implemented physically [15,16], though currently this approach is not experimentally viable.

The 'basic algorithm' that follows from the definition of a separable state is simply a straightforward search for a convex combination of separable pure states that gives the required density matrix within precision δ. We refer to [17] for a complexity comparison of the 'basic algorithm' and our algorithm which shows that the asymptotic runtime of our algorithm is significantly less than that of the 'basic algorithm'. By clever use of global optimization routines, we anticipate that our algorithm will be of practical use in the case where M and N are small (and δ is not too small)[17]. Even for $M = N = 3$, there was previously no known better algorithm for the separability problem than the 'basic algorithm'. Note also that for $MN \leq 6$, where the PPT test is necessary and sufficient, our algorithm still offers its novel advantage in the experimental setting (as explained later).

Before describing our algorithm for separability, we note the following fact [14]: A state ρ is entangled if and only if there exists an entanglement witness [18] that detects it. An *entanglement witness* is any traceless operator $A \in \mathbf{H}_{M,N}$ for which there exists a state $\rho \in \mathcal{D}_{M,N}$ such that

$$\mathrm{tr}(A\sigma) < \mathrm{tr}(A\rho) \quad \forall \sigma \in \mathcal{S}_{M,N}. \tag{3}$$

This definition is slightly different from that used in the literature, however it substantially simplifies the description of the algorithm. Recalling that $\mathbf{H}_{M,N}$ is isomorphic to $\mathbf{R}^n$, the above definition implies that for entangled ρ there exists a hyperplane which separates ρ from the set of all separable states $\mathcal{S}_{M,N}$. If one defines the function

$$b_A \equiv \max_{\sigma \in \mathcal{S}_{M,N}} \mathrm{tr}(A\sigma), \tag{4}$$

then the set $\{X \in \mathbf{H}_{M,N} : \ \mathrm{tr}(AX) = b_A\}$ is one such hyperplane. We define σ_A to be any element of $\mathcal{S}_{M,N}$ such that $\mathrm{tr}(A\sigma_A) = b_A$. It suffices only to consider entanglement witnesses A such that $\mathrm{tr}(A^2) = 1$, that is, those which lie on the $(n-2)$-dimensional surface of a $\mathbf{0}$-centered unit-hypersphere in $\mathbf{H}_{M,N}$, where $\mathbf{0}$ is the origin (null operator in $\mathbf{H}_{M,N}$). For our purposes, however, it will be useful to characterize *all potential entanglement witnesses* by the corresponding $(n-1)$-dimensional *unit-hyperball*, $\mathcal{W}$, defined as

$$\mathcal{W} \equiv \{A \in \mathbf{H}_{M,N} : \mathrm{tr}(A) = 0, \mathrm{tr}(A^2) \leq 1\}. \tag{5}$$

We can now define the *(entanglement) witness problem*, a problem slightly harder than the separability problem:

Entanglement Witness Problem. *Given a density matrix $\rho \in \mathcal{D}_{M,N}$ and a precision δ, either assert*

SEPARABLE: *there exists a separable state σ such that $||\rho - \sigma|| < \delta$; or return*

$A \in \mathcal{W}$: *an operator such that $tr(A\sigma) < tr(A\rho) + \delta$ for all $\sigma \in \mathcal{S}_{M,N}$.*

The witness problem is thus to decide that ρ is almost separable, or to find an approximate entanglement witness for ρ. Note that any algorithm solving the witness problem also solves the separability problem. Our algorithm actually solves the witness problem.

Our algorithm is an iterative one, which calls a computationally expensive subroutine at each iteration. It is convenient to treat this subroutine as a black box, or oracle, when describing the algorithm's main structure. Simply define the *oracle*, $\mathcal{O}$, such that it takes an operator A, and returns $\mathcal{O}(A) \equiv \sigma_A$. Whichever way $\mathcal{O}(A)$ is computed, it suffices that the maximization in (4) is done over the *pure* separable states (and that σ_A is a pure state). Thus, the motivation behind reducing the separability problem to the oracle $\mathcal{O}$ is quite simple: we are exploiting the fact that the separable pure states, which are the extreme points of $\mathcal{S}_{M,N}$, are parameterized by k variables rather than n. Thus, from a practical point of view, the complexity of computing $\mathcal{O}(A)$ scales much better than that of either a brute-force search through all entanglement witnesses, or the 'basic algorithm'.

Before delving into the details, we give a high-level description of our algorithm. The algorithm maintains a set $\mathcal{K} \subseteq \mathcal{W}$ of operators which are potential entanglement witnesses for ρ. If (and only if) ρ is entangled, there exists a convex subset of $\mathcal{W}$, which we call $\mathcal{W}_\rho$, consisting of *all entanglement witnesses that detect* ρ. Throughout the algorithm, we have $\mathcal{W}_\rho \subset \mathcal{K}$. Initially, $\mathcal{K}$ is set equal to $\mathcal{W}$. In each iteration, the algorithm selects a *test-witness*, $A \in \mathcal{K}$, and computes $\sigma_A = \mathcal{O}(A)$. If $\mathrm{tr}(A\sigma_A) < \mathrm{tr}(A\rho) + \delta$, then the algorithm returns A; else, $\mathcal{K}$ is reduced and the next iteration begins. The algorithm keeps reducing the set $\mathcal{K}$ until it either finds an (approximate) element of $\mathcal{W}_\rho$, or it decides that $\mathcal{W}_\rho$ is empty, and therefore ρ is separable. To decide $\mathcal{W}_\rho$ is empty means that $\mathcal{K}$ is too small to contain $\mathcal{W}_\rho$. This requires having a lower bound on the size of $\mathcal{W}_\rho$. By

exploiting the role of δ in the problem definitions, such a lower bound can be derived in terms of δ and n. In what follows, we will ignore the role of δ, as it obfuscates the main idea of the algorithm.

We now describe how to reduce the set $\mathcal{K}$, that is, to discard elements of $\mathcal{K}$ that are not elements of $\mathcal{W}_\rho$. Suppose A is not in $\mathcal{W}_\rho$, but is sufficiently close to $\mathcal{W}_\rho$. Then, A, ρ, and σ_A can be used to define a half-space $\{X \in \mathbf{H}_{M,N} : \mathrm{tr}(KX) \geq 0\}$ that contains $\mathcal{W}_\rho$. Specifically, we have the following: Let W be any operator in $\mathcal{W}_\rho$ and suppose $A \notin \mathcal{W}_\rho$. If $\mathrm{tr}(WA) \geq 0$, then choosing

$$K \equiv (\rho - \sigma_A) - \frac{\mathrm{tr}(A(\rho - \sigma_A))}{\mathrm{tr}(A^2)} A \tag{6}$$

(and then normalizing K) gives $\mathrm{tr}(KA) = 0$ by construction, and it is easy to verify that $\mathrm{tr}(KW) > 0$. The idea is that, at each iteration, the test-witness A is chosen so that it is (approximately) in the center of the current $\mathcal{K}$ (relative to the Euclidean geometry). If the oracle $\mathcal{O}$ returns σ_A such that $\mathrm{tr}(A\rho) < \mathrm{tr}(A\sigma_A)$, then, as long as $\mathrm{tr}(WA) \geq 0$ for all $W \in \mathcal{W}_\rho$, equation (6) gives a *cutting plane* $\{X \in \mathbf{H}_{M,N} : \mathrm{tr}(KX) = 0\}$ that slices through A and $\mathbf{0}$. This allows us to discard the half of $\mathcal{K}$ consisting of operators X such that $\mathrm{tr}(KX) \leq 0$. Because $\mathcal{K}$ is being approximately halved at each step, the algorithm quickly either finds an entanglement witness for ρ or concludes that ρ is separable.

Our problem of determining whether the convex set $\mathcal{W}_\rho$ is empty using cutting planes is well studied in the field of convex optimization. However, because of our special requirement that $\mathrm{tr}(WA) \geq 0$ for all $W \in \mathcal{W}_\rho$, none of the existing algorithms can be applied directly. Fortunately, though, the analytic-central-section algorithm due to Atkinson and Vaidya [19] can be adapted for our purpose, giving an algorithm with the desired complexity.

We now describe the algorithm. Let I_{MN} be the maximally mixed state, which is properly contained in $S_{M,N}$ [7,9]. It is easy to verify that $\mathcal{W}_\rho$ must be contained in the half-space $\{X : \mathrm{tr}((\rho - I_{MN})X) \geq 0\}$. Let $K_1 \equiv (\rho - I_{MN})/||\rho - I_{MN}||$. Thus, straight away, $\mathcal{K}$ is reduced to the half-ball $\mathcal{W} \cap \{X : \mathrm{tr}(K_1 X) \geq 0\}$. The first test-witness to give to the oracle is $A = \rho - I_{MN}$ (which is along the center-line of the half-ball). If the oracle confirms that A detects ρ, then we are done. Otherwise, we use equation (6) to generate a cutting plane. By way of mathematical induction, assume that, at some later stage in the algorithm, $\mathcal{K}$ has been reduced to

$$\mathcal{K} = \mathcal{W} \bigcap \cap_{i=1}^{h} \{X : \mathrm{tr}(K_i X) \geq 0\}, \tag{7}$$

by the generation of h cutting planes $\{X : \mathrm{tr}(K_i X) = 0\}$, as described above. Recall that we want to choose a test-witness that is approximately in the center of $\mathcal{K}$. An easily computable candidate is the *analytic center*, C, of $\mathcal{K}$ [20], which is defined as the unique minimizer of the real convex function

$$F(X) \equiv -\sum_{i=1}^{h} \log(\mathrm{tr}(K_i X)) - \log(1 - ||X||^2), \tag{8}$$

defined for $X \in \mathcal{K}$. The relation $\nabla F(C) = 0$ gives $C = \frac{1-||C||^2}{2} \sum_{i=1}^{h} \frac{K_i}{\mathrm{tr}(K_i C)}$, which, by the inductive hypothesis, implies that $\mathrm{tr}(WC) \geq 0$ for all $W \in \mathcal{W}_p$.

Thus, $A = C$ is a suitable test-witness to give to the oracle and to use in equation (6). Full details of a robust algorithm are too numerous to include here but can be derived with the help of [19,21,20]. The important point is that the separability of a given density matrix can be decided with only $n \times \text{polylog}\,(n, 1/\delta)$ calls to the oracle.

Finally, we discuss how the algorithm may be used when only partial information about the state ρ is available. This is of particular use in an experimental setting. Let $\mathcal{B}$ be an orthonormal, Hermitian basis for $\mathbf{H}_{M,N}$. The state ρ can be written $\rho = \sum_{i=1}^{n} \rho_{X_i} X_i$, where $\rho_{X_i} \in \mathbf{R}$. Each coefficient ρ_{X_i} is simply the *expected value of* X_i, which equals $\text{tr}(X_i \rho)$. The expected values of all elements of $\mathcal{B}$ constitute complete information about ρ. Suppose we have only measured $j < n$ expected values. The algorithm can be applied in this reduced, j-dimensional space. If the algorithm finds a hyperplane separating ρ from $\mathcal{S}_{M,N}$, then ρ is entangled; otherwise ρ may be entangled or separable, as the j expected values are consistent with a separable state. As expected values are being gathered through experimental observation, they may be input to the algorithm. If the basis $\mathcal{B}$ is separable, then the entire procedure can be done when the subsystems are spatially separated with local operations and classical communication. The idea of searching for an entanglement witness in the span of operators whose expected values are known was discovered independently and applied, in a special case, to quantum cryptographic protocols in [22].

We would like to thank Carolina Moura Alves, Coralia Cartis, and Tom Stace for useful discussions. We acknowledge support from the EC under project RESQ (IST-2001-37559). LMI, BCT, and DC also acknowledge support from, respectively, CESG and NSERC; CMI; and NSERC and the University of Waterloo.

References

[1] L. Gurvits. Classical deterministic complexity of Edmonds' problem and quantum entanglement. In *Proceedings of the thirty-fifth ACM symposium on Theory of computing*, pages 10–19, New York, 2003. ACM Press.

[2] A. Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77:1413–1415, 1996.

[3] M. Horodecki and P. Horodecki. Reduction criterion of separability and limits for a class of distillation protocols. *Phys. Rev. A*, 59:4206, 1999.

[4] M. Nielsen and J. Kempe. Separable states are more disordered globally than locally. *Phys. Rev. Lett.*, 86:5184–7, 2001.

[5] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Distinguishing separable and entangled states. *Phys. Rev. Lett.*, 88:187904, 2002.

[6] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. A complete family of separability criteria, 2003. quant-ph/0308032.

[7] S. L. Braunstein, C. M. Caves, R. Jozsa, N. Linden, S. Popescu, and R. Schack. Separability of very noisy mixed states and implications for NMR quantum computing. *Phys. Rev. Lett.*, 83:1054, 1999.

[8] K. Zyczkowski, P. Horodecki, A. Sanpera, and M. Lewenstein. On the volume of the set of mixed entangled states. *Phys.Rev. A*, 58:883, 1998.

[9] L. Gurvits and H. Barnum. Largest separable balls around the maximally mixed bipartite quantum state. *Phys. Rev. A*, 66:062311, 2002.

[10] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195–200, 1964.
[11] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:881–884, 1969.
[12] O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello, and A. Sanpera. Detection of entanglement with few local measurements. *Phys. Rev. A*, 66:062305, 2002.
[13] M. Barbieri, F. De Martini, G. Di Nepi, P. Mataloni, G. M. D'Ariano, and C. Macchiavello. Experimental detection of entanglement with polarized photons, 2003. quant-ph/0307003.
[14] M. Horodecki, P. Horodecki, and R. Horodecki. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A*, 223:1–8, 1996.
[15] P. Horodecki and A. Ekert. Direct detection of quantum entanglement, 2001. quant-ph/0111064.
[16] H. Carteret. Noiseless circuits for the Peres criterion, 2003. quant-ph/0309216.
[17] L. M. Ioannou, B. C. Travaglione, D. Cheung, and A. K. Ekert. Improved algorithm for quantum separability and entanglement detection. *Phys. Rev. A*, 70:060303, 2004.
[18] B. M. Terhal. Bell inequalities and the separablity criterion. *Phys. Lett. A*, 271:319–326, 2000.
[19] David S. Atkinson and Pravin M. Vaidya. A cutting plane algorithm for convex programming that uses analytic centers. *Mathematical Programming*, 69:1–43, 1995.
[20] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, 1994.
[21] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. MPS-SIAM, Philadelphia, 2001.
[22] M. Curty, M. Lewenstein, and N. Lütkenhaus. Entanglement as precondition for secure quantum key distribution, 2003. quant-ph/0307151.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Generalised Entanglement Swapping

Anthony J. Short [a,1], Sandu Popescu [a,b], Nicolas Gisin [c]
[a] *HH Wills Physics Laboratory, Bristol University, UK*
[b] *Hewlett-Packard Laboratories, Stoke-Gifford, Bristol, UK*
[c] *Group of Applied Physics, University of Geneva, Switzerland.*

Abstract. We consider an analogue of entanglement-swapping for a set of black boxes with the most general correlations consistent with relativity, and introduce couplers to represent joint measurements. For binary-input/binary-output boxes, we find that no such analogue exists.

Keywords. Non-locality, entanglement, entanglement-swapping

1. Introduction

One of the most surprising aspects of quantum theory is its ability to yield *non-local* correlations, which cannot be explained by any local hidden-variable model [1,2], yet do not allow superluminal signalling. Here we consider an abstract system of black boxes which can have *any* no-signalling correlations, even those not attainable in quantum theory [3,4]. We then consider whether an analogue of 'entanglement-swapping' [5] is possible for such correlations, in which non-local correlations between Alice and Bob, and between Bob and Charlie, are used to generate non-local correlations between Alice and Charlie.

2. Correlated No-Signalling Boxes

Consider a general multi-partite system composed of N correlated subsystems, which can be moved about freely. We represent each subsystem by a black-box, which has an input (corresponding to the choice of which measurement to perform on that subsystem), and an output (which is the result of the chosen measurement). We will assume that only one input can be made to each box, and that the corresponding output is obtained immediately (without having to wait for messages to travel between the boxes). Furthermore, we assume that the probability of obtaining a given set of outputs $\mathbf{O} = \{O_1, \ldots, O_N\}$ from a system of boxes depends only on the inputs $\mathbf{I} = \{I_1, \ldots, I_N\}$ which are made to those boxes, and not on the timings of those inputs (which would be reference-frame dependent). The *state* of the boxes can therefore be represented completely by the conditional probability distribution $P(\mathbf{O}|\mathbf{I})$.

[1]Correspondence to: Dr A.J.Short, HH Wills Physics Laboratory, Tyndall Avenue, Bristol, BS8 1TL, UK Tel.: +44 117 928 8755; Fax: +44 117 925 5624; E-mail: tony.short@bristol.ac.uk

The requirement that the boxes cannot be used for superluminal signalling is represented by the following no-signalling condition: For all partitions of the boxes into two disjoint sets held by Alice and Bob, with corresponding inputs $\mathbf{x}$ and $\mathbf{y}$, and outputs $\mathbf{a}$ and $\mathbf{b}$ respectively,

$$\sum_{\mathbf{a}} P(\mathbf{O}|\mathbf{I}) = \sum_{\mathbf{a}} P(\mathbf{ab}|\mathbf{xy}) = P(\mathbf{b}|\mathbf{y}) \tag{1}$$

for some probability distribution[2] $P(\mathbf{b}|\mathbf{y})$. This ensures that Bob can learn nothing about the Alice's input from his output alone.

An example of an $N = 2$ state with binary inputs and outputs (in which Alice and Bob have one box each) [3,4], is given by

$$P_{nl}(ab|xy) = \begin{cases} 1/2 & : \quad a + b = xy \pmod 2 \\ 0 & : \quad \text{otherwise} \end{cases} . \tag{2}$$

This state generates stronger non-local correlations than anything attainable in quantum theory, achieving the maximal possible value of 4 for the Clauser-Horne-Shimony-Holt (CHSH) inequality [2], compared to $2\sqrt{2}$ for any quantum state (the Cirel'son bound), or 2 for any state describable by local hidden-variables (we will refer to any such state as 'local', and all others as 'non-local').

3. 'Entanglement-Swapping' and Couplers.

To investigate whether an analogue of entanglement-swapping exists for box states, let Alice and Bob share the non-local state $P(ab_1|xy_1)$, and Bob and Charlie share the non-local state $P(b_2c|y_1z)$. We ask whether there is a procedure which Bob can carry out such that, for some result, Alice and Charlie share a non-local final state?

The most general procedure available to Bob using the formalism introduced so far is to apply an input to one of his two boxes, and then to use its output to generate an input for his second box. Unfortunately, after Bob has input y_1 and obtained output b_1 (with probability $P(b_1|y_1)$), the state of Alice's box will 'collapse' to a particular state[3] $P_{(y_1,b_1)}(a|x) = P(a|xb_1y_1)$ which is dependent on Bob's inputs and outputs, and similarly for Charlie. The final state of Alice's and Charlie's boxes is therefore given by $P'(ac|xz) = P_{(y_1,b_1)}(a|x)P_{(y_2,b_2)}(c|z)$, which is a local state for all $\{y_1, y_2, b_1, b_2\}$.

In order to achieve entanglement swapping in quantum theory, Bob must measure his two subsystems in an entangled basis (eg. the Bell basis for two qubits). However, so far we have restricted Bob to making individual (but conditional) measurements on his two subsystems. To allow an analogue of entangled measurements, we introduce a new device called a *coupler*, which is applied jointly to Bob's two boxes, and produces a single output b'. We group together all of the

[2] $P(\mathbf{b}|\mathbf{y})$ plays an analogous role to the reduced density matrix of Bob's subsystems in quantum theory.

[3] This is analogous to the collapse of an entangled state in quantum theory, after a measurement by one party.

boxes not held by Bob into a set with inputs $\mathbf{x}$ and outputs $\mathbf{a}$ (eg. $\mathbf{x} = \{x, z\}$ and $\mathbf{a} = \{a, c\}$ above), and consider how a coupler acts on the general state $P(\mathbf{a}\, b_1 b_2 | \mathbf{x}\, y_1 y_2)$.

If the inputs $\mathbf{x}$ are applied first, then outputs $\mathbf{a}$ will be obtained with probability $P(\mathbf{a}|\mathbf{x})$, and the state of Bob's two boxes will collapse to the state $P_{(\mathbf{x},\mathbf{a})}(b_1 b_2 | y_1 y_2)$. Assuming that the effect of the coupler depends only on the probability distribution of Bob's state (and not on what is 'inside' the boxes, or how that state was obtained), the final probability distribution for $\mathbf{a}$ and b' will be given by

$$P'(\mathbf{a}b'|\mathbf{x}) = P(\mathbf{a}|\mathbf{x})\mathcal{C}\left[P_{(\mathbf{x},\mathbf{a})}(b_1 b_2 | y_1 y_2)\right] = P(\mathbf{a}|\mathbf{x}) P'_{(\mathbf{x},\mathbf{a})}(b'), \tag{3}$$

where $\mathcal{C}$ is some function characteristic of the coupler. Note that (3) automatically prevents signalling *from* Bob to other parties by applying the coupler, as

$$\sum_{b'} P'(\mathbf{a}b'|\mathbf{x}) = \sum_{b_1 b_2} P(\mathbf{a} b_1 b_2 | \mathbf{x} y_1 y_2) = P(\mathbf{a}|\mathbf{x}). \tag{4}$$

To ensure that the other parties cannot signal *to* Bob when he is using the coupler, and to reveal more about the allowed functions $\mathcal{C}$, we consider applying the coupler to a specific class of no-signalling 3-box states given by

$$P(a b_1 b_2 | x y_1 y_2) = \begin{cases} \lambda_a P_{(0,a)}(b_1 b_2 | y_1 y_2) & : \quad x = 0 \\ \sum_{a'} \lambda_{a'} P_{(0,a')}(b_1 b_2 | y_1 y_2) & : \quad x = 1, a = 0 \\ 0 & : \quad \text{otherwise} \end{cases} \tag{5}$$

in which Bob's two boxes are correlated with a single box held by Alice. If Alice enters $x = 0$ she obtains a with probability λ_a (where $\sum_a \lambda_a = 1$), and Bob's boxes collapse to the state $P_{(0,a')}(b_1 b_2 | y_1 y_2)$. However, if Alice enters $x = 1$, she obtains $a = 0$ with certainty, and Bob's boxes collapse to a mixed state. Applying the coupler to Bob's boxes, we find that the final state $P'(ab'|x)$ is no-signalling from Alice to Bob only when

$$\sum_a \lambda_a \mathcal{C}[P_{(0,a)}(b_1 b_2 | y_1 y_2)] = \mathcal{C}[\sum_{a'} \lambda'_a P_{(0,a')}(b_1 b_2 | y_1 y_2)] = P'(b'). \tag{6}$$

In order to satisfy this relation for all choices of λ_a and $P_{(0,a)}(b_1 b_2 | y_1 y_2)$, $\mathcal{C}$ must be a *linear* function of Bob's two-box probability distribution. The most general form is given by

$$P'(b') = \sum_{b_1 b_2 y_1 y_2} \xi(b', b_1 b_2 y_1 y_2) P(b_1 b_2 | y_1 y_2) + \eta(b') \tag{7}$$

$$= \sum_{b_1 b_2 y_1 y_2} \chi(b', b_1 b_2 y_1 y_2) P(b_1 b_2 | y_1 y_2), \tag{8}$$

where $\chi(b', b_1 b_2 y_1 y_2) = \xi(b', b_1 b_2 y_1 y_2) + \delta_{y_1 0}\delta_{y_2 0}\, \eta(b')$. Although we have only shown that such a coupler is no-signalling for the particular class of initial states given by (5), it is easy to verify that it will be no-signalling when applied to any initial state. The only constraint on $\chi(b', b_1 b_2 y_1 y_2)$ is that $P'(b')$ must be an allowed probability distribution (satisfying $P'(b') > 0$ and $\sum_{b'} P'(b') = 1$) for all initial two-box states $P(b_1 b_2 | y_1 y_2)$.

4. Binary-Input/Binary-Output Couplers

We now consider the particular case in which Bob's two boxes and the coupler $\mathcal{C}$ have binary inputs and outputs. The set of all bipartite binary-input/binary-output no-signalling states has been studied by Barrett et. al [4], who found that they form a 9-dimensional convex polytope with 24 vertices[4].

From (8), it is easy to see that the non-negativity and normalisation of $P'(b')$ will be maintained for convex combinations of initial states, so it is sufficient to check that $P'(b')$ is non-negative and normalised for the 24 extremal states. As convex combinations of couplers are also valid couplers, the allowed distributions $\chi(b', b_1 b_2 y_1 y_2)$ form a convex polytope bounded by the hyperplanes corresponding to the non-negativity conditions. This coupler polytope was found to be 9-dimensional, with 82 vertices[5]. Surprisingly, all 82 extremal couplers can be simulated by Bob simply using his original box inputs and outputs (eg. Bob inputs $y_1 = y_2 = 0$ and outputs $b' = b_1 b_2$). This means that there are no non-trivial couplers which could correspond to an entangled measurement, and therefore no possibility of Bob introducing non-local correlations between Alice and Charlie using a coupler. This result extends to the general case in which the coupler output b' is not binary, as any coupler which generated non-local correlations between Alice and Charlie for some output $b' = \beta$ could be used to simulate a binary output coupler with the same effect (i.e. $\chi'(0, b_1 b_2 y_1 y_2) = \chi(\beta, b_1 b_2 y_1 y_2)$, and $\chi'(1, b_1 b_2 y_1 y_2) = \sum_{b' \neq \beta} \chi(b', b_1 b_2 y_1 y_2)$), which we have proven is impossible. There is therefore no analogue of entanglement-swapping for binary-input/binary-output box states, showing that while generalised box correlations may be stronger than those in quantum theory, they are harder to manipulate. This offers a possible explanation why these stronger correlations are not realised in nature.

Acknowledgements

The authors would like to thank Serge Massar, Lluis Masanes, Tobias Osbourne and Jonathan Barrett for interesting discussions, and acknowledge support from the E.P.S.R.C (IRC "Quantum Information Processing") of the U.K., and from the E.U. under European Commission project RESQ (contract IST-2001-37559).

References

[1] J.S.Bell, *Physics* **1**, 195 (1964).
[2] J.F.Clauser, M.A.Horne, A.Shimony, and R.A.Holt, *Phys. Rev. Lett.* **23** 880 (1969).
[3] S.Popescu and D.Rohrlich, *Found. Phys.* **24**, 379 (1994).
[4] J.Barrett, N.Linden, S.Massar, S.Pironio, S.Popescu, and D.Roberts, *Phys. Rev. A* **71**, 022101 (2004).
[5] M.Żukowski, A.Zeilinger, M.A.Horne, and A.K.Ekert, *Phys. Rev. Lett.* **71**, 4287 (1993).

[4] Up to local reversible operations, 8 of these vertex-states are equivalent to the state $P_{nl}(ab|xy)$, and the remaining 16 are equivalent to the local state $P(b_1 b_2|y_1 y_2) = \delta_{b_1 0}\, \delta_{b_2 0}$

[5] The coupler polytope was obtained using the polytope-solver LRS, by D.Avis.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Quantum Information Processing with Low-Dimensional Systems

Alexander Yu. Vlasov
Federal Radiology Center (IRH), 197101, Mira Street 8, St.-Petersburg, Russia

Abstract. A 'register' in quantum information processing is a composition of k quantum systems, 'qu*d*its'. The dimensions of Hilbert spaces for one qu*d*it and whole quantum register are d and d^k respectively, but we should have the possibility of preparing arbitrary *entangled* state of these k systems. Preparation and arbitrary transformations of states are possible with a universal set of quantum gates and for any d, this universal set may consist of gates acting only on single systems and neighbouring pairs. Here, we revisit methods of construction of Hamiltonians for such universal sets of gates and as a concrete new example, we consider the case with qutrits. Quantum tomography is also revisited briefly.

Keywords. qubit, qutrit, qudit, universality, quantum tomography

1. Introduction

Discrete quantum variables are a basic resource in quantum computing. A *qubit* is described by a two-dimensional Hilbert space and systems with higher dimensions are also widely used [1].

Quantum mechanics with *continuous variables* may be more intuitive due to the correspondence principle. For example, after change of classical momentum q and coordinate p to quantum operators $\hat{q}$, $\hat{p}$ in simple Hamiltonians we almost directly produce the correct quantum description.

On the other hand, it is impossible to introduce the $\hat{p}$, $\hat{q}$ operators for a system with a finite-dimensional Hilbert space. Even if for large dimensions $d \gg 2$ the continuous case could be used as an approximate model of a discrete system, it does not seem possible for low dimensions.

In 1928, Weyl suggested a method of quantization, appropriate both for the finite and infinite-dimensional cases [2]. The basic idea is to use instead of operators of coordinate $\hat{q}$ and momentum $\hat{p}$ their exponents with pure imaginary multipliers and instead of Heisenberg commutation relations to write the *Weyl system*

$$\hat{U} = e^{i\alpha\hat{p}}, \quad \hat{V} = e^{i\beta\hat{q}}, \quad \hat{U}\hat{V} = e^{i\alpha\beta}\hat{V}\hat{U}. \tag{1}$$

An analogue of the Weyl-Heisenberg commutation relations Eq. (1) may be written also for discrete quantum variables such as qubits. This is reviewed in

Sec. 2 Eqs. (3,4). Due to the relevance of the considered scheme for the finite-dimensional case with spin-1/2 systems (so-called *Jordan-Wigner representation*) Weyl wrote:

> "Because of these results I feel certain that the general scheme of quantum kinematics formulated above is correct. But the field of discrete groups offers many possibilities which we have not as yet been able to realize in Nature; " ...

Nowadays, due to many applications of the Weyl pair, Eq. (3), in quantum computation, error correction, cryptography and tomography the note about *many possibilities in field of discrete groups* looks quite justified.

In quantum information processing many *entangled* systems are put to use. In Sec. 3 we consider specific constructions with the tensor product of Weyl matrices. Due to the regular algebraic structure, it is convenient to use such operators for construction of *nonlocal* Hamiltonians for universal sets of quantum gates in any dimension. In Sec. 4, we present methods of construction of the sets for any $d \geq 2$ together with an example of Hamiltonians for qutrits.

Quantum tomography describes effective measurement procedures for quantum systems and ensembles. Weyl pairs are also useful tool in this area. These are discussed briefly in Sec. 5.

2. Pauli and Weyl Matrices

The Pauli matrices $\hat{\sigma}_x = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $\hat{\sigma}_y = \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$, $\hat{\sigma}_z = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, which satisfy the anti-commutation relations

$$\hat{\sigma}_\nu \hat{\sigma}_\mu + \hat{\sigma}_\nu \hat{\sigma}_\mu = 2\delta_{\mu\nu}, \tag{2}$$

may be generalized for $d > 2$ using a *Weyl pair*, *i.e.*, two $d \times d$ matrices [2]

$$\hat{U} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \hat{V} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \zeta & 0 & \cdots & 0 \\ 0 & 0 & \zeta^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \zeta^{d-1} \end{pmatrix} \tag{3}$$

with the property that

$$\hat{U}\hat{V} = \zeta \hat{V}\hat{U}, \quad \zeta^d = 1, \quad \zeta = e^{2\pi i/d}. \tag{4}$$

In quantum information processing the matrices Eq. (3) were widely used at first in the theory of quantum error correcting codes [1,3].

There are different ways to introduce analogues of *three* Pauli matrices, *e.g.*,

$$\hat{X} = \hat{U}, \quad \hat{Y} = \zeta^{(d-1)/2} \hat{U}\hat{V}, \quad \hat{Z} = \hat{V}. \tag{5}$$

$$\hat{X}^d = \hat{Y}^d = \hat{Z}^d = \hat{1}, \quad \hat{X}\hat{Y} = \zeta \hat{Y}\hat{X}, \quad \hat{Y}\hat{Z} = \zeta \hat{Z}\hat{Y}, \quad \underline{\hat{X}\hat{Z} = \zeta \hat{Z}\hat{X}}. \tag{6}$$

For $d = 2$ with $\zeta = \zeta^{-1} = -1$ we have $\hat{X} = \hat{\sigma}_x$, $\hat{Y} = \hat{\sigma}_y$, $\hat{Z} = \hat{\sigma}_z$, and Eq. (6) is reduced to Eq. (2). For $d > 2$ and $\zeta \neq \zeta^{-1}$ it is necessary to remember about an *order* (*e.g.*, $\hat{Z}\hat{X} = \zeta^{-1}\hat{X}\hat{Z}$).

3. Systems With n Qu$d$its

The Hilbert space of a system with n qu$d$its ($d = \dim \mathcal{H}_d \geq 2$) is a tensor product with n terms $\mathcal{H}_d^{\otimes n} = \underbrace{\mathcal{H}_d \otimes \cdots \otimes \mathcal{H}_d}_{n}$. Let us introduce a family with $2n$ operators

$$d = 2: \quad \hat{\mathfrak{x}}_{2k-1} = \hat{\sigma}_z^{\otimes(k-1)} \otimes \hat{\sigma}_x \otimes \hat{1}^{\otimes(n-k)}, \quad \hat{\mathfrak{x}}_{2k} = \hat{\sigma}_z^{\otimes(k-1)} \otimes \hat{\sigma}_y \otimes \hat{1}^{\otimes(n-k)}. \quad (7)$$

$$d > 2: \quad \hat{\mathfrak{x}}_{2k-1} = \hat{Z}^{\otimes(k-1)} \otimes \hat{X} \otimes \hat{1}^{\otimes(n-k)}, \quad \hat{\mathfrak{x}}_{2k} = \hat{Z}^{\otimes(k-1)} \otimes \hat{Y} \otimes \hat{1}^{\otimes(n-k)}. \quad (8)$$

For any given dimension $d \geq 2$, the operators $\hat{\mathfrak{x}}_k$ (for $k = 1, \ldots, 2n$) have the properties

$$\hat{\mathfrak{x}}_j^d = \hat{1}, \quad \hat{\mathfrak{x}}_j \hat{\mathfrak{x}}_k = \zeta \hat{\mathfrak{x}}_k \hat{\mathfrak{x}}_j, \quad j < k, \quad \zeta = e^{2\pi i/d}. \quad (9)$$

$$(a_1\hat{\mathfrak{x}}_1 + a_2\hat{\mathfrak{x}}_2 + \cdots + a_{2n}\hat{\mathfrak{x}}_{2n})^d = (a_1^d + a_2^d + \cdots + a_{2n}^d)\hat{1}. \quad (10)$$

For $d = 2$ Eqs. (9,10) define the generators of the *Clifford algebra* [4]

$$\hat{\mathfrak{x}}_j \hat{\mathfrak{x}}_k + \hat{\mathfrak{x}}_k \hat{\mathfrak{x}}_j = 2\delta_{jk}, \quad (a_1\hat{\mathfrak{x}}_1 + \cdots + a_{2n}\hat{\mathfrak{x}}_{2n})^2 = a_1^2 + \cdots + a_{2n}^2. \quad (11)$$

For the generalized case of $d > 2$, Eq. (9) defines an *algebra of the quantum plane* A_ζ^{2n} [5].

4. Universality

The elements described in Sec. 2,3 enable the construction of Hamiltonians for a universal set of quantum gates with simple methods of decomposition and useful properties:

1. It is a set of one- and two-qu$d$it gates (a gate for given $\hat{H}$ is $\hat{G}^\tau = e^{-i\hat{H}\tau}$).
2. Two-qu$d$it gates act on pairs of *neighbouring* systems (qu$d$its).
3. Hamiltonians of the two-qu$d$it gates are *diagonal*.

The basic idea [6,7,8] is to start with elements $\hat{\mathfrak{x}}_k \hat{\mathfrak{x}}_{k+1}^\dagger$ and use them for the construction of Hamiltonians of one- and two-qu$d$it gates. In proofs of universality, it is typical to use elements generated via commutators [9,10], but due to Eq. (9) these always have the property that $[\hat{A}, \hat{B}] = (1 - \zeta^l)\hat{A}\hat{B}$ with an integer l and it produces some simplification.

It is useful also to exchange $\hat{X} \leftrightarrow \hat{Z}^\dagger$ in Eq. (8) and define elements

$$\hat{\mathfrak{z}}_{2k-1}^\dagger = \hat{X}^{\otimes(k-1)} \otimes \hat{Z} \otimes \hat{1}^{\otimes(n-k)}, \quad \hat{\mathfrak{z}}_{2k}^\dagger = \hat{X}^{\otimes(k-1)} \otimes \hat{Y} \otimes \hat{1}^{\otimes(n-k)} \quad (12)$$

to make two-qu$d$it operators, like $\hat{Z}_k^\dagger \hat{Z}_{k+1}$ in Eq. (13) below, diagonal

$$\hat{\mathfrak{z}}_{2k-1}\hat{\mathfrak{z}}_{2k}^\dagger = \hat{X}_k, \quad \hat{\mathfrak{z}}_{2k}\hat{\mathfrak{z}}_{2k+1}^\dagger = \hat{Z}_k^\dagger \hat{Z}_{k+1}. \quad (13)$$

where we have used the more compact notation $\hat{X}_k \equiv \hat{1}^{\otimes(k-1)} \otimes \hat{X} \otimes \hat{1}^{\otimes(n-k)}$, *etc.*

• *Qubits.* $\hat{X} \equiv \hat{\sigma}_x$, $\hat{Z} \equiv \hat{\sigma}_z$. The elements Eq. (13) are Hermitian and may be used as Hamiltonians. The Hamiltonians generate only a *subgroup* of $\mathrm{SU}(2^n)$ and this subgroup is isomorphic with $\mathrm{Spin}(2n)$ [6], *i.e.*, has only quadratic dimension. It demonstrates an important class of *nonuniversal* gates and has analogues both in optical realizations [11] and in "fermionic" implementations [12].

Theorem I *Hamiltonians $\hat{X}_k$ of one-qubit gates together with diagonal Hamiltonians $\hat{Z}_k\hat{Z}_{k+1}$ of two-qubit gates are not universal: they generate only qudratic subgroup of $\mathrm{SU}(2^n)$ isomorphic to $\mathrm{Spin}(2n)$.*

It is enough for universality to add two Hamiltonians of one-qubit gates [6]

$$\hat{\mathfrak{z}}_1 = \hat{Z}_1, \quad \hat{\mathfrak{z}}_1\hat{\mathfrak{z}}_2\hat{\mathfrak{z}}_3 = \hat{Z}_2. \tag{14}$$

Theorem II *Hamiltonians $\hat{X}_k$, $\hat{Z}_1$, $\hat{Z}_2$ of one-qubit gates together with diagonal Hamiltonians $\hat{Z}_k\hat{Z}_{k+1}$ of two-qubit gates generate a universal set of quantum gates.*

• *Qu-dits.* For $d > 2$, $\hat{\mathfrak{z}}_k$ and $\hat{\mathfrak{z}}_k\hat{\mathfrak{z}}_{k+1}^\dagger$ are not Hermitian, but it is enough to split each term into two Hermitian parts [7].

Theorem III *Hamiltonians $\hat{Z}_1+\hat{Z}_1^\dagger$, $i(\hat{Z}_1-\hat{Z}_1^\dagger)$, $\hat{X}_k+\hat{X}_k^\dagger$, $i\big(\hat{X}_k-\hat{X}_k^\dagger\big)$ of one-qudit gates together with diagonal Hamiltonians $\hat{Z}_k\hat{Z}_{k+1}^\dagger+\hat{Z}_{k+1}\hat{Z}_k^\dagger$, $i(\hat{Z}_k\hat{Z}_{k+1}^\dagger-\hat{Z}_{k+1}\hat{Z}_k^\dagger)$ of two-qudit gates generate a universal set of quantum gates in $\mathrm{SU}(d^n)$.*

• *Qutrits.* Let us consider the Hamiltonians for the simlest case of a qutrit. Initial (non-Hermitian) matrices here are

$$\hat{X} = \hat{U} = \begin{pmatrix} 0&1&0\\0&0&1\\1&0&0\end{pmatrix}, \quad \hat{Z} = \hat{V} = \begin{pmatrix} 1&0&0\\0&\omega&0\\0&0&\bar{\omega}\end{pmatrix}, \quad \omega = e^{2\pi i/3}. \tag{15}$$

Let us construct a universal set of quantum gates using the elements from Th. III and their linear combinations. The Hamiltonians for one-qutrit gates are thus

$$\begin{pmatrix} 1&0&0\\0&-1&0\\0&0&0\end{pmatrix}, \ \begin{pmatrix} 1&0&0\\0&0&0\\0&0&-1\end{pmatrix}, \ \begin{pmatrix} 0&1&1\\1&0&1\\1&1&0\end{pmatrix}, \ \begin{pmatrix} 0&i&-i\\-i&0&i\\i&-i&0\end{pmatrix}. \tag{16}$$

A single Hamiltonian, $\hat{H}_d$, is sufficient to generate the two required two-qutrit interaction in Th. III

$$\hat{H}_d = |0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1| + |2\rangle\langle 2| \otimes |2\rangle\langle 2| - \tfrac{1}{3}\hat{1} \otimes \hat{1}. \tag{17}$$

5. Quantum Tomography

It is useful also to recollect briefly the methods of quantum tomography related with the operators introduced above. Assume that we have an unlimited source of quantum systems with *unknown* state described by a density matrix $\hat{\rho}$. A simple set of measurement devices may be described by projectors $\hat{\mathsf{P}}_k = |\phi_k\rangle\langle\phi_k|$. Such devices produce a "click" with probability

$$p_k = \mathrm{Tr}(\hat{\mathsf{P}}_k \hat{\rho}) = \langle \phi_k | \hat{\rho} | \phi_k \rangle. \tag{18}$$

Which sets of vectors $|\phi_k\rangle \in \mathcal{H}_d$ are necessary for complete reconstruction of any density matrix, if all probabilities p_k in Eq. (18) are estimated after a sufficiently large series of measurements? In general, a density matrix may be described by $d^2 - 1$ real parameters and it corresponds to the minimal number of such vectors.

If d is a power of a prime number $d = p^m$, special symmetric sets based on $d+1$ *mutually unbiased bases* (MUB) exist [13]. Such terminology is used because for any two vectors in different bases, $|\langle \phi \mid \varphi \rangle|^2 = 1/d$. The construction for power $m > 1$ intensively uses theory of *Galois fields* [13], but if d itself is prime, there is quite a visual model [14]:

Theorem IV *If dimension d is prime number, the eigenvectors of $d+1$ matrices $\hat{Z}, \hat{X}, \hat{X}\hat{Z}, \ldots, \hat{X}\hat{Z}^{d-1}$ produce a MUB.*

The eigenvectors of matrix $\hat{Z}$ are simply the computational basis δ_{kl} and the d^2 eigenvectors of the other d matrices $\hat{X}\hat{Z}^n$ have components $\phi_k = \frac{1}{\sqrt{d}} e^{2\pi i (ak^2 + bk)/d}$, *i.e.*, each such vector is described by two fixed numbers $a, b = 0, \ldots, d-1$.

For a qutrit, this corresponds to four matrices $\hat{Z}, \hat{X}, \hat{X}\hat{Z}, \hat{X}\hat{Z}^2$ with 12 eigenvectors

$$\left(\begin{array}{c|c|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right), \frac{1}{\sqrt{3}} \left(\begin{array}{c|c|c} 1 & 1 & 1 \\ 1 & \omega & \bar{\omega} \\ 1 & \bar{\omega} & \omega \end{array}\right), \frac{1}{\sqrt{3}} \left(\begin{array}{c|c|c} 1 & 1 & 1 \\ \omega & \bar{\omega} & 1 \\ \omega & 1 & \bar{\omega} \end{array}\right), \frac{1}{\sqrt{3}} \left(\begin{array}{c|c|c} 1 & 1 & 1 \\ \bar{\omega} & 1 & \omega \\ \bar{\omega} & \omega & 1 \end{array}\right). \tag{19}$$

For tomography of an arbitrary system, it is always possible to use a representation of the Hilbert space as a tensor product of prime dimensions, but at least for powers of primes such a procedure is not optimal [13,14].

The MUB is yet not maximally symmetric, because the scalar product for elements in *different* bases is nonzero, but in the *same* basis all vectors are orthogonal and any scalar product is null.

The **SIC-POVM conjecture** [15] suggests the existence of other symmetric sets: *in any dimension d there exist d^2 vectors with the property $|\langle \phi \mid \varphi \rangle|^2 = 1/(d+1)$ for any two vectors and all the vectors may be produced from a single vector $|\phi\rangle \mapsto \hat{X}^a \hat{Z}^b |\phi\rangle$.*

6. Conclusion

In quantum information processing, systems with higher dimension of Hilbert space (qu*d*its) and continuous quantum variables can be used instead of qubits [16]. Any dimension may have specific properties, say for the purpose of quantum tomography it is useful to distinguish case of prime dimension ($d = 2, 3, 5, 7, 11, \ldots$), power of prime ($d = 4, 8, 9, \ldots$), and the "general" case ($d = 6, 10, 12, \ldots$).

On the other hand, there are general methods, discussed above, which work with systems in any dimension. Despite obvious differences between quantum

systems with low dimension and the continuous limit, there are some tools like Weyl pairs, that may provide useful constructions and hints in many cases.

References

[1] D. Gottesman, "Fault-tolerant quantum computation with higher-dimensional systems," *Lect. Not. Comp. Sci.* **1509** 302–313 (1999); `quant-ph/9802007`.

[2] H. Weyl, *Gruppentheorie und Quantenmechanik* (S. Hirzel Verlag, Leipzig 1928); *The theory of groups and quantum mechanics* (Dover Publications, New York 1931).

[3] E. Knill, "Non-binary unitary error bases and quantum codes," *LANL report LAUR-96-2717*, `quant-ph/9608048` (1996).

[4] J. E. Gilbert and M. A. M. Murray, *Clifford algebras and Dirac operators in harmonic analysis* (Cambridge University Press, Cambridge 1991).

[5] Yu. I. Manin, *Quantum groups and noncommutative geometry* (Les publ. CRM, Montréal 1988). C. Kassel, *Quantum groups* (Springer–Verlag, New York 1995).

[6] A. Yu. Vlasov, "Clifford algebras and universal sets of quantum gates," *Phys. Rev. A* **63** 054302 (2001); `quant-ph/0010071`.

[7] A. Yu. Vlasov, "Noncommutative tori and universal sets of non-binary quantum gates," *Jour. Math. Phys.* **43** 2959–2964 (2002); `quant-ph/0012009`.

[8] A. Yu. Vlasov, "Algebra of quantum computations with higher dimensional systems," *Proc. SPIE* **5128** 29–36 (2003); `quant-ph/0210049`.

[9] D. Deutsch, A. Barenco, and A. Ekert, "Universality in quantum computation," *Proc. R. Soc. London Ser. A* **449**, 669–677 (1995).

[10] D. P. DiVincenzo, "Two-bit gates are universal for quantum computation," *Phys. Rev. A* **51**, 1015–1022 (1995).

[11] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," *Nature* **409** 46–52 (2001); `quant-ph/0006088`.

[12] B. M. Terhal and D. P. DiVincenzo, "Classical simulation of noninteracting-fermion quantum circuits," *Phys. Rev. A* **65** 032325 (2002); `quant-ph/0108010`.

[13] W. K. Wootters, "Picturing qubits in phase space," *IBM J. Res. Dev.* **48** 99–110 (2004); `quant-ph/0306135` [and references therein].

[14] A. Yu. Vlasov, "On symmetric sets of projectors (for reconstruction of a density matrix)," in V. Dorofeev, Yu. Pavlov, and E. Poberii (eds), *Gravitation, Cosmology and Elementary Particles* (SPb UEF Publishing, St.-Petersburg 2004) 147–154; `quant-ph/0302064`.

[15] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, "Symmetric informationally complete quantum measurements," *Jour. Math. Phys.* **45**, 2171–2180 (2004); `quant-ph/0310075`.

[16] S. Lloyd and S. L. Braunstein, "Quantum computation over continuous variables," *Phys. Rev. Lett.* **82**, 1784–1787 (1999); `quant-ph/9810082`.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Local Information and Nonorthogonal States

Jonathan Walgate [1]
Institute for Quantum Information Science, University of Calgary

Abstract. Measurements of quantum systems extract only classical information. The resulting dichotomy between physically real and physically discoverable information underlies many of the most interesting phenomena of quantum theory. One such phenomenon is the capacity of quantum systems to encode information globally such that exhaustive measurement of their parts might not reveal it; this is the problem of local distinguishability. Here we focus on this problem and study simple qubit-based multipartite systems. We prove three results. First, arbitrarily large sets of arbitrarily multipartite states can always be found such that the states are completely mutually nonorthogonal, yet measurements on just one copy suffice to reduce the set of possible states of a system to two. Second, all locally indistinguishable sets of quantum states can be rendered perfectly distinguishable by the addition of just one system in a complementary set of purely nonorthogonal states. Third, sets of orthogonal product states exist that are perfectly LOCC distinguishable, yet only the completely nonorthogonal subsystem must always be measured during successful LOCC protocols.

Keywords. distinguishability, local information, nonorthogonality, trine states.

All sets of orthogonal states can be perfectly distinguished by a projection measurement. Sets of multipartite states, however, often cannot be perfectly distinguished by local operations and classical communication; this holds true even of product states [1]. Here we examine sets of multipartite states that can be perfectly distinguished using LOCC, but only via complicated local protocols bearing little resemblance to a global projection. This lets us prove three simple but perhaps counterintuitive results [2].

First we show that even faced with LOCC constraints, information can be encoded in sets of nonorthogonal, arbitrarily multipartite qubit states such that it may be extracted unambiguously and with certainty. It is impossible for the state of such a system to be perfectly identified, of course; our measurements do the very next best thing - they reduce the number of possible states of the system to two.

Theorem 1 *Let n and m be any positive integers. We can construct a set of n completely nonorthogonal product states $\{|\phi_i\rangle\}$, shared between m parties, with*

[1]Correspondence to: Jonathan Walgate, Institute for Quantum Information Science, University of Calgary, Calgary, Alberta, T2N 1N4, Canada. Tel.: +1 403 210 8493; Fax: +1 403 289 3331; E-mail: jon.walgate@qubit.org.

the following property: When a system S is prepared in one of those n states, the m parties, knowing only this fact, can use local operations and classical communication to determine with certainty a pair of states in one of which the system must have been prepared.

Proof: Construct the set of n states as follows. Let the states be written:

$$|\phi_i\rangle = \bigotimes_{j=1}^{m} |\theta_i^j\rangle \qquad where \qquad |\theta_i^j\rangle = \bigotimes_{k=1}^{l} |\omega_{ik}^j\rangle. \tag{1}$$

Here $|\theta_i^j\rangle$ is the state occupied by the jth party's subsystem. Each subsystem is composed of unentangled qubits in states $|\omega_{ik}^j\rangle$. All the states $|\omega_{ik}^j\rangle$ will be one of three possibilities:

$$|\triangle_1\rangle = |0\rangle, \quad |\triangle_2\rangle = \cos\frac{\pi}{3}|0\rangle + \sin\frac{\pi}{3}|1\rangle, \quad |\triangle_3\rangle = \cos\frac{\pi}{3}|0\rangle - \sin\frac{\pi}{3}|1\rangle. \tag{2}$$

Since $\langle\triangle_i|\triangle_j\rangle > 0$ for all i, j, this guarantees that $\langle\phi_i|\phi_j\rangle > 0$ for all i, j so our overall set of states is completely nonorthogonal.

The three states $|\triangle_1\rangle$, $|\triangle_2\rangle$, and $|\triangle_3\rangle$ will be assigned to the qubits $|\omega_{ik}^j\rangle$ so that they satisfy the following constraint: for every triplet of states that are members of the set $\{\phi_i\} : \{|\phi_x\rangle, |\phi_y\rangle, |\phi_z\rangle\}$ with $x < y < z$, there must exist at least one qubit such that $|\omega_{xk}^j\rangle = |\triangle_1\rangle$, $|\omega_{yk}^j\rangle = |\triangle_2\rangle$ and $|\omega_{zk}^j\rangle = |\triangle_3\rangle$. In other words, every subset of three of the overall states has some qubit associated with it occupying a different possible state in each of the three cases.

There is no limit to the number of qubits we employ, so such an assignment is always possible. There are ${}^nC_3$ triplets associated with the n-member set $\{|\phi_i\rangle\}$, so trivially ${}^nC_3$ qubits will suffice. In practise many fewer qubits will be needed for $n > 3$, as a single qubit can satisfy the above assignment for many different triplets of possible states. It does not matter how we allot the individual qubits to the parties sharing the states $\{|\phi_i\rangle\}$, as we will be measuring them separately. One party might own all the qubits, or there might be as many parties as there are qubits. There is no upper limit to the number of qubits. Thus m can be any number we choose. There may be many possible state assignments satisfying the constraint; it does not matter which we select.

Now that we have constructed a set of n nonorthogonal product states shared between m parties, we can perform the following measurement procedure. Every qubit is individually measured in the same way, according the following three POVM elements:

$$T_1^\dagger T_1 = \frac{2}{3}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad T_2^\dagger T_2 = \frac{2}{3}\begin{pmatrix} \frac{3}{4} & -\frac{\sqrt{3}}{4} \\ -\frac{\sqrt{3}}{4} & \frac{1}{4} \end{pmatrix}, \quad T_3^\dagger T_3 = \frac{2}{3}\begin{pmatrix} \frac{3}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{1}{4} \end{pmatrix}. \tag{3}$$

It is easy to verify that $\langle\triangle_i|T_i^\dagger T_i|\triangle_i\rangle = 0$, so each outcome eliminates one of the three possible qubit states. This measurement is performed by all parties separately upon every qubit in their possession. By the above construction every

triplet of the overall states $\{|\phi_i\rangle\}$ has a corresponding qubit in states (2). Thus for *every* triplet of states, one member of that triplet will have been eliminated as a possibility by some qubit measurement. When all the results are in, this means exactly two members of the set $\{|\phi_i\rangle\}$ will remain as possible states. If three had remained, they would comprise a triplet, one member of which would have been eliminated, so three possibilities did not remain. □

We are used to measurements on nonorthogonal systems yielding unreliable information with certainty - simple optimal measurements do this. We are also used to measurements on nonorthogonal systems yielding reliable information with some probability - simple conclusive measurements do this. However, it is unusual to find a measurement of nonorthogonal systems that yields so much reliable information with certainty. This seems somehow to conflict with the role nonorthogonality notionally plays in quantum theory. Nevertheless, such measurements exist, and they are fundamental to the task of extracting information from unentangled systems.

This result directly implies the next: no set of orthogonal quantum states is so profoundly locally immeasurable, however entangled it may be, that it does not become perfectly locally distinguishable with the addition of just one adjoint system in a commensurate set of nonorthogonal states. This is counterintuitive; the state of the adjoint system is *globally* unidentifiable. How then can it render the most locally indistinguishable set of states, of a system with which it is not dynamically connected, not just *more* distinguishable, but *perfectly* locally distinguishable?

Theorem 2 *Let* $\{|\psi_i\rangle_{BC...}\}$ *be a set of n orthogonal multipartite quantum states, shared between Bob, Carol, and so on. There always exists an extended set of states* $\{|\phi_i\rangle_A \otimes |\psi_i\rangle_{BC...}\}$ *which is locally distinguishable, and where the set of states* $\{|\phi_i\rangle_A\}$ *in Alice's possession are all mutually nonorthogonal:* $\forall i, j \quad \langle\phi_i|\phi_j\rangle \neq 0$.

Proof: This follows directly from Theorem 1, and the fact that two orthogonal multipartite quantum states, no matter how entangled, can always be locally distinguished [3]. We can choose Alice's adjoint states $\{|\phi_i\rangle_A\}$ to be constructed according to the protocol outlined above, with $m = 1$. Then however locally indistinguishable the set of states $\{|\psi_i\rangle_{BC...}\}$ is, the joint states $\{|\phi_i\rangle_A \otimes |\psi_i\rangle_{BC...}\}$ can be LOCC distinguished as follows: Alice performs measurement (3) upon all her qubits, and eliminates all but two possible states without disturbing the subsystems of the other parties. She communicates her result, whereupon Bob, Carol *et al.* need only locally distinguish between the two remaining orthogonal alternatives. This is always possible [3]. □

Our final theorem establishes another curious result: nonorthogonal subsystems are not only helpful, and occasionally vital, in locally identifying orthogonal states - sometimes the only locally vital subsystem is the nonorthogonal one!

Theorem 3 *Let the states* $\{|\psi_i\rangle\}$ *be defined by equation (4), where* $|\triangle_i\rangle_A$ *are defined as in equation (2).* $\{|\psi_i\rangle\}$ *is not perfectly distinguishable using local projection measurements and classical communication.* $\{|\psi_i\rangle\}$ *is locally distinguishable, however using POV measurements. Given one run of an LOCC protocol identifying the state of the system with certainty, only one party must necessarily have*

performed a measurement and that party is Alice, whose states bear none of the orthogonality of $\{|\psi_i\rangle\}$.

$$|\psi_1\rangle = |\triangle_1\rangle_A|0\rangle_B|0\rangle_C|0\rangle_D, \quad |\psi_2\rangle = |\triangle_2\rangle_A|1\rangle_B|+\rangle_C|+\rangle_D,$$
$$|\psi_3\rangle = |\triangle_3\rangle_A|+\rangle_B|1\rangle_C|-\rangle_D, \quad |\psi_4\rangle = |\triangle_1\rangle_A|-\rangle_B|-\rangle_C|1\rangle_D. \tag{4}$$

Proof: Let the four partial states shared by Bob, Carol and David be written $\{|\phi_i\rangle_{BCD}\}$. $\{|\phi_i\rangle_{BCD}\}$ are known to be indistinguishable using LOCC; they are the simplest set of states exhibiting nonlocality without entanglement [1]. *Any* nontrivial local measurement by Bob, Carol or David has some measurement outcome that renders at least one pair of the four states nonorthogonal. Because no pair of the overall product states $\{|\psi_i\rangle\}$ is orthogonal in Alice's Hilbert space, this property extends to the four-partite states as well. Alice must make the first measurement. Yet all possible projection measurements by Alice stand a nonzero chance of failing to eliminate any of her four possible states. After such a result, the system would only be known to be in the locally indistinguishable set of states $\{|A\rangle_A \otimes |\phi_i\rangle_{BCD}\}$. Thus $\{|\psi_i\rangle\}$ is not perfectly distinguishable using local projection measurements and classical communication.

Let Alice instead performs measurement (3). In line with the three possible results, she will eliminate either states $|\psi_1\rangle$ and $|\psi_4\rangle$, state $|\psi_2\rangle$, or state $|\psi_3\rangle$. This leaves Bob, Carol and David holding three possible states drawn from the set $\{|\phi_i\rangle_{BCD}\}$. Any three members of this state are clearly locally distinguishable. In the case of Alice's first measurement result, eliminating $|\psi_1\rangle$ and $|\psi_4\rangle$, a single measurement by David will identify the actual state, and neither Bob nor Carol will have been required to act by the LOCC protocol. In the second case, Carol can perform a $\{|0\rangle\langle 0|, |1\rangle\langle 1|$ measurement, and depending on its result either Bob or David will be in a position to distinguish the remaining two states. Thus it is not guaranteed that David will be required to act by the LOCC protocol. Therefore the only party whose measurement must necessarily be performed while locally distinguishing the states $\{|\psi_i\rangle\}$ is Alice, since her partners are forbidden from acting first [4]. □

When tasked to distinguish orthogonal quantum states globally, we would rightly dismiss nonorthogonal subsystems as irrelevant. Using LOCC, they are essential. The constraints of locality can be deeply counterintuitive.

I would like to thank the Alberta Ingenuity Fund and the Pacific Institute for the Mathematical Sciences for funding this research. I would also like to thank Somshubhro Bandyopadhyay, Lucien Hardy, Barry Sanders and Tony Short for their helpful comments and insights.

References

[1] C.H. Bennett, D.P. DiVincenzo, C.A. Fuchs, T. Mor, E. Rains, P.W. Shor, J.A. Smolin, and W.K. Wootters, Phys. Rev. A **59**, 1070 (1999)
[2] J. Walgate, paper in preparation.
[3] J. Walgate, A.J. Short, L. Hardy, and V. Vedral, Phys. Rev. Lett. **85**, 4972 (2000)
[4] J. Walgate and L. Hardy, Phys. Rev. Lett. **89**, 147901 (2002)

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Optimal Alphabets for Noise-Resistant Quantum Cryptography

Denis Sych [a,1], Boris Grishanin [a] and Victor Zadkov [a]
[a] *International Laser Center and Department of Physics*
M. V. Lomonosov Moscow State University

Abstract. Possibilities of increasing the critical error rate of quantum key distribution (QKD) protocols are investigated. We consider QKD protocols with discrete alphabets, letters of which form regular polyhedrons on the Bloch sphere (tetrahedron, octahedron, cube, icosahedron, and dodecahedron, which have 4, 6, 8, 12, and 20 vertices respectively) and a QKD protocol with continuous alphabet, which corresponds to the limiting case of a polyhedron with infinite number of vertices. The stability of such QKD protocols against intercept–resend and optimal eavesdropping attacks on the individual information carriers is studied in detail. It is shown that all these QKD protocols have approximately the same critical error rates. In the case of optimal eavesdropping strategies, after basis reconciliation, the QKD protocol with continuous alphabet surpasses all other protocols in terms of noise-resistance. Without basis reconciliation the protocol with the highest critical error rate has a tetrahedron-type alphabet. Additional increase of the dimensionality of the quantum alphabet leads to a further increase of the critical error rate.

Keywords. quantum key distribution, quantum cryptography, critical error rate, quantum alphabet

1. Introduction

Since the 1980s, when the idea of quantum cryptography was first proposed [1], a number of different quantum key distribution (QKD) protocols have been suggested [2,3,4,5]. Despite their diversity, all of them are based on a beautiful idea employing a basic "no-cloning" principle of quantum mechanics—the impossibility of copying arbitrary quantum states [6]. Thanks to this, an eavesdropper cannot intercept the quantum communication channel without disturbing a transmitting message, providing it contains a set of *incompatible*, i.e. essentially quantum, states. Moreover, any attempt to obtain any information about this set of states inevitably disturbs the transmitted message.

Keeping this advantage of quantum physics for cryptography in mind, any QKD protocol uses messages entirely composed of an incompatible set of quantum states or so called *quantum alphabet* that consists of incompatible "letters". Various QKD protocols are distinguished in essence only by different alphabets, which ensure secure message transmission up to a critical error rate that determines the protocol efficiency.

[1]Correspondence to: Denis Sych, Moscow 119992, Leninskie Gory, Khokhlova ul., 1, International Laser Center MSU; Tel.: +7 095 939 5173; Fax: +7 095 939 3113; E-mail: denis.sych@gmail.com.

All known QKD protocols using carriers of information with a finite-dimensional Hilbert space are based on discrete quantum alphabets, i.e. with a fixed number of letters. The first QKD protocol proposed in 1984 by Bennett and Brassard (BB84) [1], gives an example of so-called "prepare and measure" protocol in which four incompatible quantum states, forming two mutually nonorthogonal bases, are used. The alphabet of the six-state protocol [4] is composed of three mutually nonorthogonal bases of the two-dimensional Hilbert space, which makes this protocol totally symmetrical on the Bloch sphere and leads to the fact that the critical error rate of the six-state protocol is higher than that of the BB84-protocol [4,7]. In two-dimensional case, there is a commonly accepted opinion that the six-state protocol has the best efficacy among all other one-way QKD protocols. Further increasing the critical error rate is basically connected with increasing the dimension of the Hilbert space of the quantum alphabet [8,9,10].

In this paper we will clarify whether varying of the quantum alphabet in the Hilbert space of fixed dimension could improve the QKD protocol efficacy or not. In other words, can we surpass the six-state protocol efficacy, even in the two-dimensional case, only by the varying the set of letters in the alphabet or not?

In order to answer this question, we introduce QKD protocols with discrete alphabets, letters of which form regular polyhedrons on the Bloch sphere (tetrahedron, octahedron, cube, icosahedron, and dodecahedron, which have 4, 6, 8, 12, and 20 vertices) [11], and a QKD protocol with a continuous alphabet, which corresponds to the limiting case of a polyhedron with infinitive number of vertices [12]. The six-state protocol corresponds to the octahedron–type alphabet, so by analogy with it we will call such protocols 4-, 6-, 8-, 12-, 20-, and ∞-state protocols, respectively. All these protocols imply standard steps like those of the BB84 (transmission of a raw key, basis reconciliation, etc.). Efficacy of these protocols can be also calculated in the same way as for other standard QKD protocols, i.e. with the help of information analysis based on calculation of the mutual Shannon information between different two-partite subsystems of the tripartite system Alice–Eve–Bob [13]. At the same time, the ∞-state protocol has some specific properties which we outline below.

From a theoretical point of view, the key difference in analysis of the ∞-state protocol with respect to QKD protocols with discrete alphabets lies in calculation of the amount of information that can be encoded with the help of continuous alphabet. A natural quantitative measure for the amount of information is the standard mutual Shannon information functional of the classical input–output (Alice-Bob) joint probability distribution P_{AB}, and for the ∞-state protocol it transforms to the so-called non-selected compatible information [14,15]

$$I_{AB} = \iint\limits_{\alpha\,\beta} P_{AB}(d\alpha, d\beta) \log_2 \frac{P_{AB}(d\alpha, d\beta)}{P_A(d\alpha) P_B(d\beta)}$$

with joint probability distribution $P_{AB}(d\alpha, d\beta) = \mathrm{Tr}_{\mathrm{AB}}\big(\big[\hat{E}_A(d\alpha) \otimes \hat{E}_B(d\beta)\big]\hat{\rho}_{AB}\big)$, where a set of positive operators representing a continuous non-orthogonal expansion of the unit operator (or the continuous POVM) is $\hat{E}_{A,B}(d\nu) = |\nu\rangle_{A,B}\langle\nu|_{A,B}\, dV_\nu$, and $\mathrm{d}V_\nu$ is the volume differential normalized to the dimension of Hilbert space D.

From a practical point of view, the most significant difference between a QKD protocol with a continuous alphabet and QKD protocols with discrete alphabets lies in the basis reconciliation procedure. In contrast to a discrete alphabet, one cannot perform an exact basis reconciliation procedure for the ∞-states protocol, because one needs to

transmit an infinite amount of information about a point from a continuum. Therefore, we suggest the use of an *approximate* basis reconciliation procedure for the ∞-states protocol [12].

One more detail of the ∞-states protocol, which follows directly from the approximate basis reconciliation procedure, is how to estimate the level of Eve's interference. One of the most accepted in the literature is the quantum bit error rate (QBER) $Q = 1 - N/N_{\max}$, where N is the number of correctly transmitted letters and $N_{\max}$ is the total number of transmitted letters. In order to resolve some contradictions with the use of QBER in case of ∞–state protocol (see also Ref. [16]), we suggest that the fidelity of data transmission be defined not as a relative number of correctly transferred letters, but as the relative amount of correctly transferred information $\tilde{Q} = 1 - I/I_{\max}$, where I is the amount of information per one message with the presence of eavesdropping and $I_{\max}$ is its maximal possible value without eavesdropping. We will call this measure, by analogy with QBER, the *mutual information error rate* or MIER.

2. Comparison of the QKD Protocols

For determining the critical error rate up to which the QKD protocol ensures the perfect secrecy of the transmitted data, one needs generally to prove the unconditional security of the QKD protocol [17,18]. However, in this paper we are not going to carry out an unconditional security proof, but perform a *comparative* analysis of the protocols under different attacks by Eve. We will examine two key strategies of eavesdropping—intercept–resend and optimal eavesdropping.

The simplest eavesdropping strategy is the intercept–resend strategy, when Eve measures a message transmitted over a quantum channel in an arbitrary orthogonal basis and then transmits to Bob the results of this measurement. It is clear that by using such a strategy, Eve knows exactly the information received by Bob and therefore secure data transmission between Alice and Bob is impossible. Therefore, the maximal possible level of errors which can be corrected, while ensuring that the transmission is a secure one, cannot exceed the level of errors Q caused by the intercept–resend strategy of eavesdropping. As a result, calculation of the error rate due to this strategy of eavesdropping gives the upper bound of the protocol efficacy for any applied strategy of eavesdropping.

Calculating Q for the considered QKD protocols, we get $Q_{4-\text{state}} = Q_{6-\text{state}} = Q_{\infty-\text{state}} = 1/3 \simeq 0.333$ and $Q_{12-\text{state}} = Q_{20-\text{state}} = 74/225 \simeq 0.329$. It is interesting to note that despite similar geometrical symmetry of all discussed alphabets, they have, nevertheless, different error rates: 4-, 6- and ∞-state protocols have equal error rates but these exceed those of the 12- and 20-state protocols.

Considering the optimal eavesdropping strategy, we note that it has been proved that in one–way communication schemes, when only Alice can send qubits to Bob, a secure connection between Alice and Bob is possible if the amount of information Bob received from Alice exceeds information Eve received either from Alice or Bob [19]. This condition can be written as $I_{AB} > \max(I_{AE}, I_{BE})$.

We will call Eve's eavesdropping strategy *optimal* if Eve extracts from the transmitting message the maximum information at the given level of interference which causes the respective level of errors, which can be written as $I_{AE,BE} = \max\limits_{I_{AB}=\text{const}} I_{AE,BE}$.

Table 1. Critical error rate $\tilde{Q}_0$ for the 4-, 6-, 8-, 12-, 20-, and ∞-state QKD protocols.

Number of letters	4	6	8	12	20	∞
Before (without) basis reconciliation	0.650	0.630	0.607	0.597	0.589	0.600
After (with) basis reconciliation	*)	0.806	0.805	0.804	0.805	0.811

*) The standard basis reconciliation procedure cannot be explicitly introduced for the 4–state protocol, as far as its alphabet cannot be represented as a combination of orthogonal bases.

The results of calculation of critical error rates are summarized in Table 1, where we can see that without basis reconciliation the 4-state protocol has the best maximum value of the critical error rate. After the basis reconciliation procedure is applied we see, however, that the ∞-state protocol has highest critical error rate among the considered QKD protocols. In other words, the information characteristics of a six-state protocol can be surpassed even in the case of a two-dimensional Hilbert space due to a better choice of the alphabet.

Briefly discussing a potential of using multidimensional spaces for Bob and Alice spaces ($D > 2$) for improving the properties of QKD protocols, we note that for the infinite–dimensional case, the maximum possible error rate is equal to 100%, by contrast with 50% for the two-dimensional case [12]. Therefore, one can conclude that there are no principal restrictions on increasing the efficacy of the QKD protocols by increasing the dimensionality of the Hilbert space because there is no upper threshold set by the intercept-resend strategy.

Results of calculation of critical error rate for the multidimensional ∞-state protocol show a qualitatively novel property with respect to two-dimensional one—by increasing the dimensionality of the Hilbert space, the critical error rate for this protocol increases and in the limit of infinite-dimensional space the protocol becomes non-threshold.

References

[1] Ch.H. Bennett and G. Brassard, *Proceedings of the IEEE International Conference on Computer, System and Signal Processing, Bangalore, India*, p. 175, IEEE, New York, 1984.
[2] A.K. Ekert, *Phys. Rev. A* **67** (1991), 661.
[3] Ch.H. Bennett, *Phys. Rev. Lett.* **68** (1992), 3121.
[4] D. Bruss, *Phys. Rev. Lett.* **81** (1998), 3018.
[5] F. Grosshans and P. Grangier, *Phys. Rev. Lett.* **88** (2002), 057902.
[6] W.K. Wootters and W.H. Zurek, *Nature (London)* **299** (1982), 802.
[7] H. Bechmann-Pasquinucci and N. Gisin, *Phys. Rev. A* **59** (1999), 4238.
[8] H. Bechmann-Pasquinucci and W. Tittel, *Phys. Rev. A* **61** (2000), 062308.
[9] M. Bourennane, A. Karlsson, and G. Bjork, *Phys. Rev. A* **64** (2001), 012306.
[10] N.J. Cerf, M. Bourennane, A. Karlsson, and N. Gisin, *Phys. Rev. Lett.* **88** (2002), 127902.
[11] D.V. Sych, B.A. Grishanin and V.N. Zadkov, *Sov. J. Quantum. Electron.* **35**(1) (2005), 80.
[12] D.V. Sych, B.A. Grishanin and V.N. Zadkov, *Phys. Rev. A* **70** (2004), 052331.
[13] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, *Rev. Mod. Phys.* **74** (2002), 145.
[14] B.A. Grishanin, *Probl. Inf. Transm.* **38** (2002), 26.
[15] B.A. Grishanin and V.N. Zadkov, *J. Commun. Technol. Electron.* **47**(9) (2002), 933.
[16] C.A. Fuchs and A. Peres, *Phys. Rev. A* **53** (1996), 2038.

[17] H.-K. Lo and H.F. Chau, *Science* **283** (1999), 2050.
[18] P.W. Shor and J. Preskill, *Phys. Rev. Lett.* **85** (2000), 441.
[19] I. Csiszar, and J. Koerner, IEEE Trans. Inform. Theor., **24** (1978), 339.

This page intentionally left blank

Chapter 2

Quantum Algorithms and Error Correction

This page intentionally left blank

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Quantum Algorithms and Complexity

Michele Mosca [a,b,1]
[a] *Institute for Quantum Computing*
Department of Combinatorics & Optimization, University of Waterloo and St. Jerome's University, Waterloo, Canada
[b] *Perimeter Institute for Theoretical Physics, Waterloo, Canada*

Keywords. Quantum Algorithms. Quantum Complexity.

1. Introduction and Overview

This article corresponds to three lectures I gave at the NATO summer school. I will not reiterate everything from these lectures, as most of the material already appears in the literature. Rather, I will mostly overview the topics I chose to cover in the lectures, including some pointers to the literature.

I had several goals when preparing and delivering these lectures. Firstly, I wanted to convey to the participants the significance of quantum algorithms. For this, I included a brief summary of computational complexity, and throughout I discuss the conjectured or proved complexity advantages that quantum algorithms provide.

However I wanted these lectures to be more than a review or overview of quantum algorithms. I wanted to teach the interested participants the "tools of the trade", as much as is reasonable in such a short time. Thus I tried to give an understanding of how the algorithms work and present much of the technical know-how behind quantum algorithms. I also tried to sketch the past, present and speculate about the future of quantum algorithms.

This article is not intended to be a comprehensive history or survey of quantum algorithms and complexity.

2. A Brief History of Quantum Algorithms

In 1982, Richard Feynman [30] delivered a lecture on the difficulty of simulating quantum mechanical systems. He speculated that some sort of "quantum" computer, with components that evolved according to the laws of quantum theory, could in a very natural and efficient way simulate such quantum mechanical systems.

[1]Supported by Canada Research Chair, CIAR, CFI, OIT, ORDCF, NSERC, MITACS and ARDA. mmosca@iqc.uwaterloo.ca

This line of thought was developed in more detail by Lloyd [47], Zalka [62] and others. To a large extent, this family of quantum algorithms hasn't been the focus of much attention by quantum computer scientists. However, I hope this changes, since I think the combination of expertise in quantum algorithms and of expertise in the problems related to the behaviour of quantum systems (including classical algorithms for them) is very fertile territory for the development of intrinsically new quantum algorithmic techniques and for solving important problems in physics.

In his 1985 paper, David Deutsch [25] set out to explicitly place the famous Church-Turing thesis in a physical context. In this paper he also details the first quantum algorithms. It solves the problem of deciding whether a function $f : \{0,1\} \to \{0,1\}$ satisfies $f(0) = f(1)$ or $f(0) \neq f(1)$, given a reversible black-box that evaluates f. Any classical algorithm must evaluate f twice, but a quantum algorithm can solve the problem with only one evaluation.

Deutsch's algorithm was generalized further by Deutsch and Jozsa [26], Bernstein and Vazirani [13], and Simon [59]. A crucial breakthrough occurred in 1994 when Shor, inspired by the techniques Simon used, found an efficient algorithm for integer factorization and for finding discrete logarithms [57]. Shor's algorithms inspired the generalizations of Boneh and Lipton [14], Grigoriev [34] and Kitaev [44].

Not long after, it was realized that these problems were special cases of the "hidden subgroup" problem. Kitaev's algorithm looked qualitatively different, but it was also eventually realized that Kitaev's and Shor's approach to solving these problems were in some sense equivalent [19].

In 1996, Lov Grover [35] developed a quantum algorithm for searching for a solution to $f(x) = 1$ given a means (e.g. a "black-box") to evaluate f. This algorithm was nicely analyzed and generalized by Boyer, Brassard, Høyer and Tapp into a technique known as "amplitude amplification" [12]. Brassard, Høyer and Tapp also realized that one could combine quantum searching with the period-finding tools used for factoring, and developed quantum counting [17,16].

I will stop this brief history and overview for now, and carry on with more technical discussions that will give the essence behind most of the algorithms I have just discussed. After these detailed discussions, I will conclude with a brief discussion of other quantum algorithms and algorithmic paradigms.

3. Introduction to Computational Complexity

The "computational complexity" of an algorithm is some measure of the resources used by the algorithm. We typically measure resources like time, space, or elementary operations. We typically represent the complexity of an algorithm as a function of the size of the input to the algorithm. For example, a simple algorithm for multiplying two n-bit numbers might take $O(n^2)$ steps. We also typically consider the complexity of an algorithm on a "worst-case" input of size n. So it will take at most $O(n^2)$ steps to multiply *any* two n-digit numbers.

The computational complexity of a problem is a measure of the intrinsic difficulty of the problem and corresponds to the computational complexity of the

"best" algorithm that solves the problem. We know the computational complexity of multiplication (in terms of elementary logic gates) is at most $O(n^2)$ since the elementary algorithm for multiplication takes that many steps. However this is only an upper bound. Better algorithms can multiply using $O(n \log n \log \log n)$ steps [56]. Any algorithm requires at least $\Omega(n)$ steps, since any algorithm must look at each bit of the inputs.

In general, one can find upper bounds on the computational complexity of a problem by finding an algorithm that solves the problem and providing an upper bound for the complexity of that algorithm. Finding non-trivial lower bounds on the computational complexity of a problem has proved to be a much harder problem.

We hope that the more general framework of quantum computation will help us find non-trivial lower bounds and some new relationships between the difficulty of various problems. There as been some progress in this direction (e.g. [43,3,1]), but it has still proved to be a very difficult task to find non-trivial lower bounds for quantum computers as well. The limitations of certain algorithmic approaches and of certain proof techniques can be determined by proving lower bounds in the "black-box model" (e.g. see [10], [20], [9], [6] for more on the black-model and the lower bound techniques).

For convenience, let us restrict attention to "decision" problems, where the answer is either YES or NO. Decision problems can be treated as the problem of recognizing elements of a *language*. This framework of language recognition problems might seem awkward at first, but much of computational complexity theory has been developed using this terminology, so it is useful to be somewhat familiar with it. Appendix A explains this formalism in a bit more detail and defines a few of the most common complexity classes one will encounter in the quantum computing literature.

The classes defined in the appendix are P (which corresponds to problems efficiently solvable on a deterministic classical computer), BPP (which corresponds to problems efficiently solvable with high probability of correctness on a probabilistic classical computer), BQP (which corresponds to problems efficiently solvable with high probability of correctness on a quantum computer), NP (which corresponds to problems for which a YES answer has a succinct certificate that can be checked efficiently on a deterministic classical computer), and $PSPACE$ (which corresponds to the set of problems that can be solved using a polynomial amount of space).

Figure 1 illustrates the known relationships between the complexity classes we just defined. For example, clearly $P \subset BPP \subset BQP \subset PSPACE$, and $P \subset NP \subset PSPACE$. Unfortunately, to date, there is no proof that any of the containments drawn are strict. But it is widely believed that $P \neq NP$, and that $NP \neq PSPACE$. We also expect that $BPP \neq BQP$. At present, it would not be surprising if $BPP = P$ [39]. It is interesting to note that resolving the question of whether $P = NP$ is considered one of the greatest open problems in mathematics (e.g. it is one of the million dollar "Millenium Problems" of the Clay Mathematics Institute http://www.claymath.org/millennium/P_ vs _ NP/).

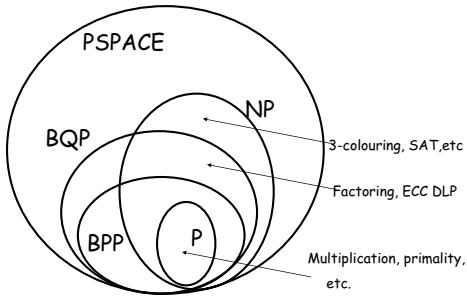


Figure 1. This diagram illustrated the known relationships between some of the most important complexity classes. At present, none of the inclusions are known to be strict. For example, there is currently no proof that $P \neq PSPACE$.

I have only sketched a very small number of complexity classes. See, for example, Scott Aaronson's complexity "zoo" http://www.complexityzoo.com for a listing of virtually all the complexity classes studied to date.

The main challenge of quantum algorithmics is to find problems that are in BQP but not in BPP. The study of these complexity classes and the relationships between them can be helpful in understanding the power of algorithms. The black-box lower bound methods provide some evidence that BQP does not contain all of NP (see e.g. [61]), though the possibility is not explicitly ruled out at present. The lower bound methods also prove the optimality of several of the black-box quantum algorithms we describe below.

The next section of this tutorial article will focus on the development of quantum algorithms.

4. Basics of Quantum Algorithms

We can restrict attention to the circuit model of quantum computation, where our information is stored in a string of qubits, which have some naturally defined basis whose elements we label $|0\rangle$ and $|1\rangle$. Each qubit can be measured with respect to this "computational basis".

Quantum theory tells us that we can *in principle* measure a string of n qubits with respect to any orthonormal basis of n-qubit states, $\mathcal{B} = \{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_{2^n}\rangle\}$. We call this a *Von Neumann measurement* with respect to basis $\mathcal{B}$. Such a measurement applied to a superposition $|\Psi\rangle = \sum_x \alpha_x |\psi_x\rangle$ should output a label 'x' with probability $|\alpha_x^2|$ and leave the system in state $|\psi_x\rangle$. How can such a measurement be implemented, at least approximately, in practice? It suffices to be able to implement, with a quantum circuit, the unitary transformation U that maps $|\psi_j\rangle \mapsto |j\rangle$ (where j refers to a string of bits representing the integer j). The transformation U would map $|\Psi\rangle = \sum_x \alpha_x |\psi_x\rangle \mapsto \sum_x \alpha_x |x\rangle$. A measurement in the computational basis would output label 'x' with probability $|\alpha_x^2|$ and leave the system in state $|x\rangle$. Applying $U^\dagger$ then gives us the state $|\psi_x\rangle$. We have thus applied a measurement with respect to the basis $\mathcal{B}$. In summary, given the means to measure in the computational basis, we can implement a Von

Neumann measurement with respect to basis $\mathcal{B}$ by conjugating the computational basis measurement by a unitary operation that performs a basis change from $\mathcal{B}$ to the computational basis. Viewing an operator conjugated by some basis change is used widely in quantum computation.

The simple Hadamard gate is the operator that changes the $\{|0\rangle, |1\rangle\}$ basis into the $\{|0\rangle + |1\rangle, |0\rangle - |1\rangle\}$ basis, and vice versa. A convenient way of summarizing the action of the Hadamard gate is that it maps $|b\rangle \mapsto |0\rangle + (-1)^b|1\rangle$, $b \in \{0,1\}$, and since it is self-inverse, it also maps $|0\rangle + (-1)^b|1\rangle \mapsto |b\rangle$.

Imagine someone gave you the qubit $|0\rangle + (-1)^b|1\rangle$ and challenged you to find b. Measuring the qubit in the computational basis would give you a random bit. However the Hadamard gate in some sense decodes the information 'b' encoded in the relative phase into something measurable in the computational basis.

Consider the controlled-NOT gate (or CNOT), usually described as the transformation that maps $|0\rangle|b\rangle \mapsto |0\rangle|b\rangle$ and $|1\rangle|b\rangle \mapsto |1\rangle NOT|b\rangle$, where the NOT gate (also called the X gate) maps $|0\rangle \mapsto |1\rangle$ and $|1\rangle \mapsto |0\rangle$. The CNOT can be represented by the matrix

$$\begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{pmatrix}.$$

One often says that the CNOT 'does nothing' if the first register is in state $|0\rangle$ and 'flips the second bit' if the first register is in state $|1\rangle$. This correctly describes its behaviour on computational basis states, in which case, it appears that 'nothing happens' to the first qubit (the 'control' bit) and any changes occurs on the second qubit (the 'target' bit). Its behaviour on arbitrary two qubit states is the linear extension of its behaviour in the computational basis. This notion that nothing happens to the control bit can be misleading since something does indeed happen to the control bit when the CNOT is applied to more general states.

For example, the CNOT maps

$$\begin{aligned} |0\rangle(|0\rangle + |1\rangle) &\mapsto |0\rangle(|0\rangle + |1\rangle) \\ |0\rangle(|0\rangle - |1\rangle) &\mapsto |0\rangle(|0\rangle - |1\rangle) \\ |1\rangle(|0\rangle + |1\rangle) &\mapsto |1\rangle(|0\rangle + |1\rangle) \\ |1\rangle(|0\rangle - |1\rangle) &\mapsto -|1\rangle(|0\rangle - |1\rangle). \end{aligned}$$

In other words, conjugating the CNOT by $I \otimes H$ gives us

$$\begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&-1 \end{pmatrix}.$$

So we can say that the CNOT applies a relative phase shift of -1 with respect to the basis $\{|0\rangle + |1\rangle, |0\rangle - |1\rangle\}$ in the target register when the control register is in state $|1\rangle$.

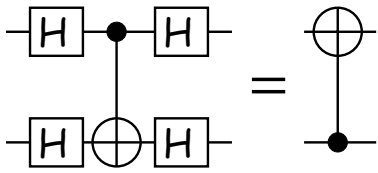


Figure 2. Conjugating the CNOT by Hadamard gates swaps the role of the 'control' and 'target' bits.

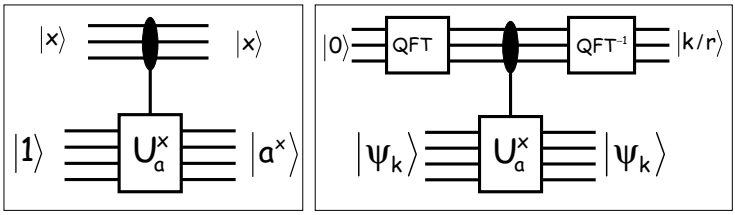


Figure 3. The figure on the left is a simple exponentiation circuit, that multiplies the values stored in the second register by a^x when x is stored in the first register. On the right we conjugate the exponentiation circuit by $QFT \otimes I$. When we consider the action of this operation with an eigenstate of $|\,y\rangle \mapsto |\,ya^x\rangle$ in the second register, we see that information about reciprocals of the period r of a is encoded in the first register.

Note that the CNOT also maps

$$(|\,0\rangle + |\,1\rangle)(|\,0\rangle + |\,1\rangle) \mapsto (|\,0\rangle + |\,1\rangle)(|\,0\rangle + |\,1\rangle)$$
$$(|\,0\rangle - |\,1\rangle)(|\,0\rangle + |\,1\rangle) \mapsto (|\,0\rangle - |\,1\rangle)(|\,0\rangle + |\,1\rangle)$$
$$(|\,0\rangle + |\,1\rangle)(|\,0\rangle - |\,1\rangle) \mapsto (|\,0\rangle - |\,1\rangle)(|\,0\rangle - |\,1\rangle)$$
$$(|\,0\rangle - |\,1\rangle)(|\,0\rangle - |\,1\rangle) \mapsto (|\,0\rangle - |\,1\rangle)(|\,0\rangle - |\,1\rangle)$$

In this basis, the CNOT seems to flip the state of the first register when the second register is in state $|\,0\rangle - |\,1\rangle$, and do 'nothing' to the second register. In other words, conjugating the CNOT by $H \otimes H$ gives us a CNOT with the roles of the target and control bit swapped, as illustrated in Figure 2.

To summarize, although in the computational basis we appear to be doing something to the second register controlled on the value of a first register, in a different basis, interesting things might be going on in the first register.

The essence of many of the most powerful algorithms can be viewed as applying a 'classical' operation with an input and output register, but in a conjugated basis. In this conjugated basis this 'classical' operation has the effect of encoding valuable information in the control register.

For example, the order-finding algorithm (a special case of period-finding that is used in the quantum factoring algorithm) is essentially looking at the classical operation for exponentiating in the basis formed by conjugating the operation by $QFT \otimes I$ (see Figure 3). The QFT transformation is the quantum Fourier transform, a natural generalization of the Hadamard gate.

A simple survey of many these early quantum algorithms can be found in [19] or also in the textbooks [52] or [41].

An elegant way of summarizing many of the problems that can be solved efficiently on a quantum computer is the Abelian Hidden Subgroup problem. Suppose we are given a black box for computing some function $f : \mathcal{G} \to X$ for some finite set X. Suppose f has the property that $f(x) = f(y)$ if and only if $x - y \in H$, where H is some subgroup of G. The hidden subgroup problem corresponding to f is to find a set of generators for the subgroup H. The algorithm for solving the hidden subgroup problem for finitely generated Abelian groups is a natural generalization of the algorithm depicted in Figure 3. The algorithm essentially conjugates a circuit for computing f by $QFT \otimes I$. In this conjugated basis, the circuit is encoding information about the subgroup H (in particular, it gives information about the orthogonal group $H^{\perp}$) in the input register of the circuit for computing f.

In the black-box model, solving the finite Abelian hidden subgroup problem on a classical computer requires a number of operations that is exponential in $\log |G|$. The quantum algorithm requires only $O(\log |G|)$ group operations and polynomial in $log|G|$ other elementary operations.

For explicit functions f (where the black-box lower bounds don't apply, since one might be able to exploit the additional information known about f), like those corresponding to order-finding in the multiplicative group of integers modulo N (which is used to factor N), or finding discrete logarithms in the multiplicative group of a finite field $GF(q)$, the best known rigorous classical algorithms require time exponential in $\sqrt{n \log n}$, where n is $\log_2 p$ and $\log_2 q$ respectively. The best-known heuristics use time exponential in $n^{\frac{1}{3}} \log^{\frac{2}{3}} n$. The quantum algorithms requires $O(n)$ group operations and $O(n^2 \log n \log \log n)$ other elementary operations.

The great success of quantum algorithms for solving the Abelian hidden subgroup problem leads to the natural question of whether it can solve the non-Abelian Hidden Subgroup problem. This question has been studied by many authors (e.g. [8,54,42,28,45,49,31,38,33]) , and quantum algorithms can be found for some non-Abelian groups, but there is no algorithm for more general non-Abelian groups, like the symmetric group (which would directly solve the graph automorphism problem e.g.). It is not clear if an efficient quantum Fourier transform for the group suffices, and several authors have studied the limitations of the natural approaches using the non-Abelian QFT.

A handful of quantum algorithms of a similar flavour have been found that seem to have superpolynomial advantage and that do not seem to be a special case of the hidden subgroup problem (e.g. [22,23,37]).

There have also been a number of algorithms offering a polynomial improvement over classical algorithms (in the "black-box model") (e.g. [35,48]). The first, and most broadly applicable, of these algorithms is Grover's quantum searching algorithm [35]. The next section describes this algorithm and some of its generalizations and applications.

5. Quantum Searching and Generalizations

Consider a function $f : \{0,1\}^n \to \{0,1\}$. Let $X_1 = \{x|f(x) = 1\}$ and $X_0 = \{x|f(x) = 0\}$. Suppose we wish to find a solution to the equation $f(x) = 1$, given the means to evaluate f. In other words, given the means to *recognize* a solution, *find* one. If we have the means to efficiently compute f, then we have the means to efficiently implement U_f which maps $|\,x\rangle \mapsto (-1)^{f(x)}|\,x\rangle$. Grover's algorithm, and its generalizations, are "black-box" algorithms, since the only way they probe f is by evaluating it.

Instead of describing the special case of quantum searching discovered by Grover, let us describe the following generalization called "amplitude amplification" [12]. Consider any algorithm A that finds a solution x to $f(x) = 1$ with probability p. Classically, one would have to repeat A on the order of $\frac{1}{p}$ times in order to boost the probability of finding a solution x to greater than $\frac{1}{2}$.

Without loss of generality [2] we can assume that A is a unitary transformation that maps $|\,00\ldots0\rangle \mapsto \sum_{x\in X_1} \alpha_x|\,x\rangle + \sum_{x\in X_0} \beta_x|\,x\rangle$. This means that the probability of measuring a "good" x (i.e. one satisfying $f(x) = 1$) is $p = \sum_{x\in X_1} |\alpha_x|^2$, and the probability of measuring a "bad" x (i.e. one satisfying $f(x) = 0$) is $1 - p = \sum_{x\in X_0} |\beta_x|^2$. We can thus rewrite $A|\,00\ldots0\rangle = \sin(\theta)|\,X_1\rangle + \cos(\theta)|\,X_0\rangle$, by setting

$$|\,X_1\rangle = \sum_{x\in X_1} \frac{\alpha_x}{\sqrt{p}}|\,x\rangle$$

and

$$|\,X_0\rangle = \sum_{x\in X_0} \frac{\beta_x}{\sqrt{1-p}}|\,x\rangle$$

where $p = \sin^2(\theta)$. If $p = 0$ or 1, we don't need to define $|\,X_1\rangle$ or $|\,X_0\rangle$, respectively.

In other words, A produces a good solution with probability amplitude $\sqrt{p}$, so one might hope that there is a way to boost the success amplitude by roughly $\sqrt{p}$ after every application of A, and thus only require $O(\sqrt{p})$ iterations of A and applications of U_f.

We define the "quantum search iterate" or "Grover iterate" to be $Q = -AU_0A^{-1}U_f$, where U_0 maps $|\,00\ldots0\rangle \to -|\,00\ldots0\rangle$ and $|\,x\rangle \mapsto |\,x\rangle$ for all $x \neq 00\ldots0$ (in other words $U_0 = I - 2|\,00\ldots0\rangle\langle 00\ldots0\,|$).

It can be verified that $QA|\,00\ldots0\rangle = \sin(3\theta)|\,X_1\rangle + \cos(3\theta)|\,X_0\rangle$. So, for example, if $p = \frac{1}{4}$, then $\theta = \frac{\pi}{6}$, and thus $\sin(3\theta) = 1$. So $QA|\,00\ldots0\rangle = |\,X_1\rangle$ and thus measuring gives a solution with probability 1.

More generally, $Q^kA|\,00\ldots0\rangle = \sin((2k+1)\theta)|\,X_1\rangle + \cos((2k+1)\theta)|\,X_0\rangle$. Thus, in order to get a high probability of measuring a good input x, one should choose k so that $(2k+1)\theta \approx \frac{\pi}{2}$, which means $k \in O(\frac{\pi}{4\theta})$. As $\theta \mapsto 0$, this corresponds to $k \in O(\frac{\pi}{4\sqrt{p}})$. Thus amplitude amplification achieves this "square-root advantage",

[2] If A is a classical algorithm, it can be turned into a fully quantum algorithm by standard methods and with modest computational overhead.

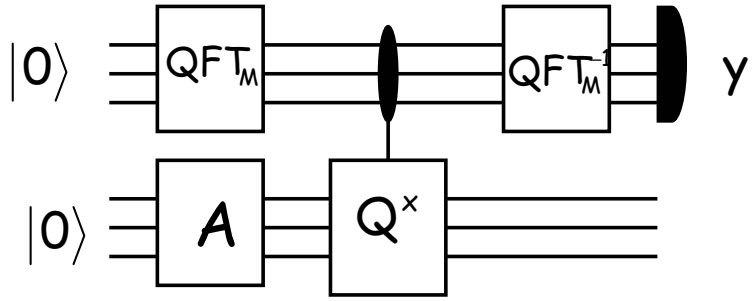


Figure 4. The output y on the control register corresponds to an eigenvalue estimate of the quantum search operator Q. With high probability the value $\sin^2(\frac{\pi y}{M})$ is a good estimate $p = \sin^2(\theta)$.

which is optimal [10]. There's a technical difficulty in the case that we don't know p, since in that case we don't know for which value of k we should apply Q^k. This problem was overcome in a simple way in [12]. We describe a related method in the next section.

One application of quantum searching is derived by observing that as we vary the parameter k the success amplitude of $Q^k A|\,00\ldots0\rangle$ is periodic with a period proportional to $\frac{1}{\theta}$. It turns out that combining the period-finding algorithm (used by Shor to factor large integers [57]) with the quantum searching algorithm gives an estimate of θ and thus of p (see [17,16]). In other words, we have an algorithm for estimating the probability of success of the algorithm A (depicted in Figure 4). If A simply uniformly samples the inputs x of f, then this amplitude estimation is approximately counting the number of solutions to $f(x) = 1$.

5.1. Searching When We Don't Know How Many Items There Are

After running the quantum counting algorithm with precision parameter M, it can be shown [50] that if one measures the target register, then the probability of obtaining a good solution x is in $\frac{1}{2} - O(\frac{1}{M\sqrt{p}})$ (let's assume for convenience that $0 < p \leq \frac{1}{2}$). One simple way to boost the probability of success [3] is to repeat this algorithm twice [4]. In this case, the probability of finding a solution is in $\frac{3}{4} - O(\frac{1}{M\sqrt{p}})$. If we try this for $M = 2, 4, 8, \ldots, 2^n$, until a solution is found (once we reach $M = 2^n$ we can do an exhaustive search), we will find a solution (if one

[3] The point of doing this is to guarantee that the probability of success p' will be strictly above $\frac{1}{r}$ once $M > \frac{1}{\sqrt{p}}$, where r is the rate at which we increase the interval size M. We want $rp' < 1$ because the expected running time will depend on the value of the geometric series $\sum_n (rp')^n$. For simplicity, we choose $r = 2$ and thus we wish to boost the probability p' above $\frac{1}{2}$. Alternatively, we could choose any rate $r < \frac{1}{2}$, as done in [12].

[4] There are other ways of boosting the probability $\frac{1}{2} - O(\frac{1}{M\sqrt{p}})$. For example, it is easy to convert the algorithm to one with probability of success $1/4 - O(\frac{1}{M\sqrt{p}})$ (it is easy to *reduce* the probability of success). For any algorithm S which outputs a solution with probability $1/4 - \epsilon$, for some small ϵ, observing after one application of the searching iterate $G = -SU_0S^{-1}U_f$ to the state $S|\,00\ldots0\rangle$ will produce a solution with probability in $1 - O(\epsilon^2)$. Grover recently came up with a clever algorithm [36] that would achieve success with probability in $1 - O(\epsilon^3)$.

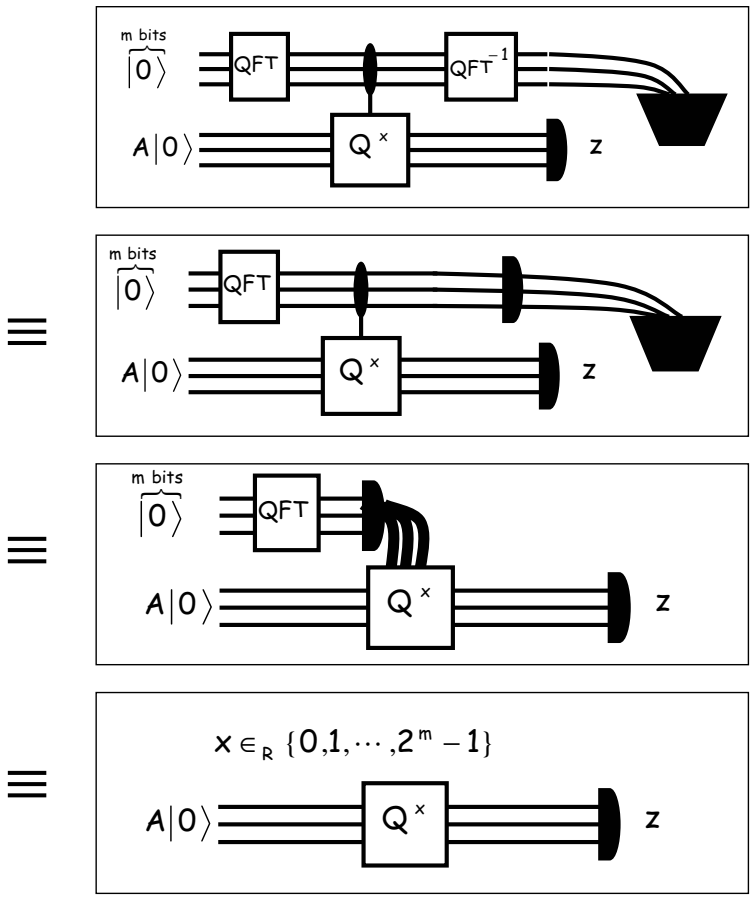


Figure 5. 1) The first picture is the quantum counting algorithm shown in Figure 4, except the control register is ignored and the target register is measured. The measured string x will be a solution to $f(x) = 1$ with probability approaching $\frac{1}{2}$ as $n \to \infty$. 2) Since the first register is discarded, we could equivalently measure the first register immediately after the controlled-Q^x operation, as depicted in the second picture. 3) The measurement on the first register commutes with the controlled-Q^x, and thus we could perform the measurement first, and then use the measured outcome, say x, control how many times to apply Q to the target register. 4) If we measure the first register before the controlled-Q^x, then it's apparent that the first register is simply acting as a random number generator that is uniformly sampling x from $\{0, 1, \ldots, 2^m - 1\}$.

exists) with a total number of applications of A and U_f that is in $O(\frac{1}{\sqrt{p}})$, thus achieving the "square-root speed-up" even when we don't know p.

This quantum algorithm for searching when the number of solutions is unknown is depicted in Figure 5. Via a sequence of simple reductions depicted in this figure, we can see that running this circuit is equivalent to running one that simply selects a random integer $x \in \{0, 1 \ldots, M - 1\}$, computes $Q^x A|\, 00 \ldots 0\rangle$, and measures. This is indeed how it is done in the original algorithm detailed in

[17].

6. Other Algorithms and Algorithmic Paradigms

There are a variety of other algorithms that offer improvements over classical algorithms. Some of these algorithms have been developed in different algorithmic paradigms.

One alternative approach for solving computational problems was inspired by the adiabatic theorem [29]. It was shown [24] that these times of adiabatic algorithms can achieve a quadratic speed-up for searching, but there is a continuing debate about the power of these adiabatic algorithms. A slightly more general type of adiabatic computation was in fact shown to be polynomially equivalent to general quantum computation [4]. Thus, this more general type of quantum adiabatic algorithm can efficiently factor large numbers or solve any other problem that a "standard" quantum algorithm can with only a polynomial overhead. The hope is that this different algorithmic paradigm will inspire the development of new quantum algorithms.

Another algorithmic approach that has been successful for classical computers has been that of random walks. This idea has been generalized in a few ways to what are called quantum walks (e.g. [5,2]) and have been used to solve various interesting problems (e.g. [7]).

The field of quantum algorithmics has experienced quite an expansion in the past five years, making a thorough survey of the field a daunting task. These lectures and notes aimed to achieve the more modest task of detailing some of the basic tools that were developed, and sketching some of the recent developments.

Let me conclude by listing some of my favourite open problems for quantum algorithms. These include the non-Abelian HSP (including Graph Automorphism and Graph Isomorphism problems), questions related to finding short vectors in a lattice (see e.g. [55]), the security of other classical cryptosystems (see e.g. [11]), NP-complete problems, and various computational problems in physics.

Appendix

A. Language Recognition Problems and Complexity Classes

Fix an alphabet, say $\Sigma = \{0, 1\}$. The set Σ^* denotes all finite length strings over that alphabet. A language L is a subset of Σ^*. In particular, L is the set of strings with some property of interest.

An algorithm solves the language recognition problem for L if it accepts any string $x \in L$ and rejects any string $x \notin L$.

For example, the problem of deciding whether an integer n (represented as a string of bits) is prime is rephrased as the problem of recognizing whether the string representing n is in the language PRIME $= \{10, 11, 010, 011, 101, 111, \ldots\}$ (which consists of the set of all strings representing prime numbers).

As another example, consider the problem of deciding whether a given graph G (represented by a string of bits) is 3-colourable is equivalent to recognizing

whether the string representing G is in the language 3−COLOURABLE, which is the set of strings representing 3−colourable graphs. [5] Note that there are only 6 possibly edges on a graph with 4 vertices (let's call them v_1, v_2, v_3, v_4), namely $e_1 = \{v_1, v_2\}, e_2 = \{v_1, v_3\}, e_3 = \{v_1, v_4\}, e_4 = \{v_2, v_3\}, e_5 = \{v_2, v_4\}, e_6 = \{v_3, v_4\}$. Thus we can naturally represent the graphs on 4 vertices by strings $x_1x_2x_3x_4x_5x_6$ of length 6, by letting $x_j = 1$ if and only if G contains the edge e_j. We can easily verify that $101111 \in$ 3−COLOURABLE since the $c(v_1) = RED$, $c(v_2) = BLUE$, $c(v_3) = RED$, $c(v_4) = GREEN$ is a valid 3−colouring.

Now that we've showed how to phrase decision problems in terms of languages, we can define various *classes* of languages. For example, we define P to be the class of languages L for which there exists a deterministic classical algorithm A running in worst-case polynomial time such that for any input $x \in \Sigma^*$ the algorithm A on input x, outputs ACCEPT if and only if $x \in L$. Note that this class does not capture the possible advantages of using randomness to efficiently solve problems.

The class BPP ("bounded-error probabilistic polynomial time") consists of all languages L for which there exists a randomized classical algorithm A running with worst-case expected polynomial time such that for any input $x \in \Sigma^*$ we have

- if $x \in L$ then the probability that A accepts x is at least $\frac{2}{3}$
- if $x \notin L$ then the probability that A accepts x is at most $\frac{1}{3}$.

It is important to note that when we say the probability that A accepts is at least $\frac{2}{3}$, this holds for any $x \in L$, and is not just an average over all $x \in L$.

It is also worth noting that there is nothing special about $\frac{2}{3}$. Any constant $\frac{1}{2} + \delta$, where $\delta > 0$ will work. For any fixed δ, we can repeat the algorithm A a total of n independent times and take the majority answer. We now get the correct answer with probability at least $1 - \epsilon^n$ for some constant ϵ, $0 < \epsilon < 1$. [51].

We traditionally view decision problems corresponding to recognizing languages in BPP as *tractable* (on a classical computer), and problems without such worst-case polynomial time solutions *intractable* (on a classical computer).

This practice of regarding BPP as capturing the informal notion of "practically efficient classical computation", or of regarding BQP as capturing the informal notion of "practically efficient quantum computation" can be controversial. For example a running time of $n^{\log\log\log n}$ is not polynomial but efficient for any practical value of n, and $(n+1)^{100}$ is a polynomial but not practical for any positive value of n. The best and quickest way to justify this practice is to follow Papadimitriou [53] and state that "adopting polynomial worst-case performance as our criterion of efficiency results in an elegant and useful theory that says something meaningful about practical computation, and would be impossible without this simplification".

Some problems seem to elude any efficient solution in the worst-case, such as deciding whether a given graph is 3-colourable. Note that although it might

[5] A graph is 3-colourable if it is possible to assign each vertex v one of three colours $c(v)$ so that any two vertices joined by an edge are coloured with different colours. Such an assignment of colours is a valid 3-colouring.

be very difficult to decide if a graph x is 3-colourable, it is very easy to check if a given colouring y is a proper 3-colouring. That is, there exists a polynomial time algorithm CHECK-3-COLOURING(a, b) such that CHECK-3-COLOURING$(x, y) = 1$ if and only if y is a valid colouring of the graph x (the algorithm simply goes through every edge of the graph x and checks that according to the colouring y the two vertices of each edge are coloured differently).

This inspires the class NP ("non-deterministic polynomial time") which consists of all languages L for which there exists a polynomial time algorithm $A(a, b)$ such that for any input $x \in \Sigma^*$ we have

- if $x \in L$ then there exists an input y such that $A(x, y)$ accepts
- if $x \notin L$ then $A(x, y)$ rejects for all y

and the length of y is bounded by a polynomial in the length of x.

There is a special subclass of problems within NP that are called "NP-complete" ([21,40,46,32]). These problems have the property that an efficient solution to any one of these problems implies the efficient solution to any problem in NP. Well-known examples of NP-complete problems include 3-colouring, the travelling salesman problem, the subset sum problem, and circuit satisfiability.

The class $PSPACE$ consists of all languages L for which there exists a classical algorithm A using worst-case polynomial space such that for any input $x \in \Sigma^*$ the algorithm A on input x, $A(x)$, accepts if and only if $x \in L$.

References

[1] S. Aaronson "Quantum Computing, Postselection, and Probabilistic Polynomial-Time", *Proceedings of the Royal Society of London A*, 461 (2005) 3473-3482.

[2] D. Aharonov, A. Ambainis, J. Kempe, U. Vazirani "Quantum Walks On Graphs", *Proceedings of ACM Symposium on Theory of Computation (STOC'01)*, (2001) 50-59.

[3] D. Aharonov, O. Regev, "Lattice Problems in NP intersect coNP", *Journal of the ACM* 52 (2005), 749-765.

[4] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, "Adiabatic Quantum Computation is Equivalent to Standard Quantum Computation", *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS'04)*, (2004) 42-51.

[5] A. Ambainis, E. Bach, A. Nayak, A. Vishwanath, J. Watrous, "One-dimensional quantum walks", *Proceedings of the 33rd ACM Symposium on Theory of Computing*, (2001) 3749.

[6] A. Ambainis, "Quantum lower bounds by quantum arguments", *J. Comput. Syst. Sci.* 64 (2002) 750-767.

[7] A. Ambainis, "Quantum walk algorithm for element distinctness", *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS'04)*, (2004) 22-31.

[8] D. Bacon, A. Childs, W. van Dam, "From optimal measurement to efficient quantum algorithms for the hidden subgroup problem over semidirect product groups", *Proc. 46th IEEE Symposium on Foundations of Computer Science (FOCS 2005)*, (2005) 469-478.

[9] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. deWolf, "Quantum lower bounds by polynomials", *Journal of the ACM*, 48 (2001) 778-797.

[10] C. Bennett, E. Bernstein, G. Brassard, U. Vazirani, *Strengths and Weaknesses of Quantum Computing*, *SIAM J. Computing*, 26 (1997) 1510 - 1523.
[11] Michael Brown, "Classical Cryptosystems in a Quantum Setting". *MMath Thesis.* University of Waterloo, 2003.
[12] Michel Boyer, Gilles Brassard, Peter Høyer, Alain Tapp, "Tight bounds on quantum searching," *Fortschritte der Physik* 56(5-5) (1998), 493-505.
[13] E. Bernstein, U. Vazirani, "Quantum Complexity Theory", *SIAM Journal on Computing Volume* 26 (1997) 1411-1473.
[14] D. Boneh, R. Lipton, "Quantum Cryptanalysis of Hidden Linear Functions (Extended Abstract)", *Proceedings of 15th Annual International Cryptology Conference (CRYPTO'95)*, (1995) 424-437.
[15] G. Brassard, P. Høyer, Alain Tapp, "Cryptology Column —Quantum Algorithm for the Collision Problem", *ACM SIGACT News*, 28 (1997) 14-19.
[16] Gilles Brassard, Peter Høyer, Michele Mosca, Alain Tapp. "Quantum Amplitude Amplification and Estimation," to appear in *Quantum Computation and Quantum Information Science, AMS Contemporary Math Series*, 2000.
[17] Gilles Brassard, Peter Høyer, Alain Tapp, "Quantum Counting," *Proceedings of the ICALP'98 Lecture notes in Computer Science* (1998), 1820-1831.
[18] H. Buhrman, C. Durr, M. Heiligman, P. Høyer, F. Magniez, M. Santha, R. de Wolf, "Quantum Algorithms for Element Distinctness", *SIAM J. Comput.*, 34 (2005) 1324-1330.
[19] Richard Cleve, Artur Ekert, Chiara Macchiavello, Michele Mosca, "Quantum Algorithms Revisited," *Proceedings of the Royal Society of London A*, 454 (1998), 339-354.
[20] R. Cleve, "An introduction to quantum complexity theory", *Collected Papers on Quantum Computation and Quantum Information Theory*, C. Macchiavello, G.M. Palma, and A. Zeilinger (Eds.) (World Scientific), (2000) 103-127.
[21] S. Cook, "The complexity of theorem proving procedures", *Annual ACM Symposium on Theory of Computing*, (1971) 151-158.
[22] W. van Dam, S. Hallgren, L. Ip, "Quantum Algorithms for some Hidden Shift Problems", *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA'03)*, (2003) 489-498.
[23] W. van Dam, S. Hallgren, "Efficient Quantum Algorithms for Shifted Quadratic Character Problems" *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*, (2003) 489-498.
[24] W. van Dam, M. Mosca, U. Vazirani, "How Powerful is Adiabatic Quantum Computation?", *Proc. 46th IEEE Symposium on Foundations of Computer Science (FOCS'01)* (2001) 279-287.
[25] D. Deutsch, "Quantum theory, the Church-Turing principle and the universal quantum computer," *Proceedings of the Royal Society of London A*, 400 (1985), 97-117.
[26] D. Deutsch, R. Jozsa, "Rapid solutions of problems by quantum computation",*Proceedings of the Royal Society of London, Series A*, 439 (1992) 553-558.
[27] C. Durr, M. Heiligman, P. Høyer, M. Mhalla, "Quantum query complexity of some graph problems", *Proc. of 31st International Colloquium on Automata, Languages, and Programming (ICALP'04)*, (2004) 481-493.
[28] M. Ettinger, P. Høyer, E. Knill, "The quantum query complexity of the hidden subgroup problem is polynomial", *Inf. Process. Lett.*, 91 (2004) 43-48.
[29] E. Farhi, J. Goldstone, S. Gutmann and M. Sipser, "Quantum Computation by Adiabatic Evolution", *eprint arXiv: quant-ph/0001106* (2000).
[30] R. Feynman, "Simulating Physics with Computers," *International Journal of Theoretical Physics*, 21(6,7) (1982), 467-488.
[31] K. Friedl, G. Ivanyos, F. Magniez, M. Santha, P. Sen, "Hidden translation and

orbit coset in quantum computing", *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing (STOC'03)*,(2003), 1-9.
[32] M. Garey and D. Johnson, *Computers and Intractability*, W. H. Freeman, (1979).
[33] M. Grigni, L. Schulman, M. Vazirani, U. Vazirani, "Quantum mechanical algorithms for the nonabelian hidden subgroup problem", *Proceedings of the thirty-third annual ACM symposium on Theory of computing (SODA'03)* (2001) 68 - 74.
[34] D. Grigoriev, "Testing Shift-Equivalence of Polynomials by Deterministic, Probabilistic and Quantum Machines", *Theor. Comput. Sci.* 180 (1997) 217-228.
[35] L. Grover, "A fast quantum mechanical algorithm for database search" *Proceedings of the 28th Annual ACM Symposium on the Theory of Computing (STOC 1996)*, (1996) 212-219.
[36] L. Grover,"Quantum Searching Amidst Uncertainty", *Unconventional Computation, 4th International Conference* (2005) 11-18.
[37] S. Hallgren, "Fast quantum algorithms for computing the unit group and class group of a number field", *Proceedings of the 37th ACM Symposium on Theory of Computing (STOC 2005)*, (2005) 468-474.
[38] S. Hallgren, M. Roetteler, P. Sen, "Limitations of Quantum Coset States for Graph Isomorphism", *eprint arXiv:quant-ph/0511148* (2005).
[39] R. Impagliazzo and A. Wigderson, "P=BPP if E requires exponential circuits: Derandomizing the XOR Lemma", *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing* (1997), 220229.
[40] R. Karp, "Reducibility among combinatorial problems", *Complexity of Computer Computations* (1972) 85-103.
[41] P. Kaye, R. Laflamme, M. Mosca "An Introduction to Quantum Computation". Oxford University Press. To appear.
[42] J. Kempe and A. Shalev, "The hidden subgroup problem and permutation group theory", *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms (SODA'05)*, (2005) 1118-1125.
[43] I. Kerenidis, R. de Wolf, "Exponential lower bound for 2-query locally decodable codes via a quantum argument", *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing (STOC 2003)*, (2003), 106-115.
[44] A. Kitaev,"Quantum measurements and the Abelian Stabilizer Problem", *Electronic Colloquium on Computational Complexity (ECCC)*, 3 (1996).
[45] G. Kuperberg, "A Subexponential-Time Quantum Algorithm for the Dihedral Hidden Subgroup Problem", *SIAM J. Comput.*, 35 (2005) 170-188.
[46] L. Levin, "Universal'nyie perebornyie zadachi" (Universal search problems: in Russian), *Problemy Peredachi Informatsii*, (1973) 9(3):265–266.
[47] S. Lloyd, "Universal Quantum Simulators", *Science*, 273 (1996) 1073 - 1078.
[48] F. Magniez, M. Santha and M. Szegedy, "Quantum algorithms for the triangle problem", *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms (SODA'05)*, (2005) 1109-1117.
[49] C. Moore, D. Rockmore, A. Russell, L. Schulman, "The power of basis selection in fourier sampling: hidden subgroup problems in affine groups", *Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms (SODA'04)*, (2004) 1113-1122.
[50] M. Mosca, "Counting by quantum eigenvalue estimation", *Theoretical Computer Science*, 264 (2001), 139-153.
[51] R. Motwani and P. Raghavan, *Randomized Algorithms* 1995, Cambridge University Press.
[52] Michael Nielson, Isaac Chuang "Quantum Computation and Quantum Information". 2000, Cambridge University Press.
[53] C. Papadimitriou, *Computational Complexity* 1994, Addison-Wesley.

[54] J. Radhakrishnan, M. Roetteler, P. Sen, "On the Power of Random Bases in Fourier Sampling: Hidden Subgroup Problem in the Heisenberg Group", *In Proceedings of the 32nd International Colloquium on Automata, Languages and Programming (ICALP)*, (2005) 1399-1411.

[55] O. Regev, "Quantum Computation and Lattice Problems", *SIAM Journal on Computing* 33 (2004) 738-760.

[56] A. Schonhage, V. Strassen, "Schnelle Multiplikation grosser Zahlen", *Computing*, 7 (1971) 281– 292.

[57] Peter Shor, "Algorithms for Quantum Computation: Discrete Logarithms and Factoring," *Proceedings of the 35th Annual Symposium on Foundations of Computer Science* (1994), 124-134.

[58] P. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer" *SIAM J. Computing*, 26 (1997) 1484 - 1509.

[59] D. Simon, "On the power of quantum computation", *Proceedings of the 35th IEEE Symposium on the Foundations of Computer Science (FOCS)*(1994) 116-123.

[60] D. Simon, "On the Power of Quantum Computation", *SIAM J. Computing*, 26 (1997) 1474-1483.

[61] U. Vazirani, "On the power of quantum computation", *Philosophical Transactions of the Royal Society of London, Series A* , 356:1759-1768, (1998).

[62] Ch. Zalka, "Efficient Simulation of Quantum Systems by Quantum Computers" *Proc. Roy. Soc. Lond. A*, 454 (1998) 313-322.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

An Introduction to Measurement Based Quantum Computation

Richard Jozsa
Department of Computer Science,
University of Bristol, Bristol BS8 1UB, UK

Abstract. In the formalism of measurement based quantum computation we start with a given fixed entangled state of many qubits and perform computation by applying a sequence of measurements to designated qubits in designated bases. The choice of basis for later measurements may depend on earlier measurement outcomes and the final result of the computation is determined from the classical data of all the measurement outcomes. This is in contrast to the more familiar gate array model in which computational steps are unitary operations, developing a large entangled state prior to some final measurements for the output. Two principal schemes of measurement based computation are teleportation quantum computation (TQC) and the so-called cluster model or one-way quantum computer (1WQC). We will describe these schemes and show how they are able to perform universal quantum computation. We will outline various possible relationships between the models which serve to clarify their workings. We will also discuss possible novel computational benefits of the measurement based models compared to the gate array model, especially issues of parallelisability of algorithms.

Keywords. quantum computation, quantum teleportation, one way quantum computer, cluster state

1. Introduction

Many of the most popular models of quantum computation are direct quantum generalisations of well known classical constructs. This includes quantum turing machines, gate arrays and walks. These models use unitary evolution as the basic mechanism of information processing and only at the end do we make measurements, converting quantum information into classical information in order to read out classical answers. In contrast to unitary evolution, measurements are irreversibly destructive, involving much loss of potential information about a quantum state's identity. Thus it is interesting, and at first sight surprising, that we can perform universal quantum computation using only measurements as computational steps [2,3,4,5,6]. These "measurement based" models are especially interesting for fundamental issues: they have no evident classical analogues and they offer a new perspective on the role of entanglement in quantum computation. They may also be interesting for experimental considerations, suggesting a

different kind of computer architecture and offering interesting possibilities for further issues such as fault tolerance [7].

We will discuss two measurement based models. Firstly we'll consider teleportation quantum computation (TQC). This is based on the idea of Gottesman and Chuang [2] of teleporting quantum gates, and was developed into a computational model by Nielsen, Leung and others [3,4]. Secondly we will consider the so-called "one way quantum computer" (1WQC) or "cluster state computation" of Raussendorf and Briegel [5,6]. Then we will discuss various possible relationships between these two models and finally consider possible computational benefits of the measurement based formalism as compared to the quantum gate array model.

The following notations will be frequently used. The Pauli operators are

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad iY = ZX = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

$\oplus$ will denote addition modulo 2. The controlled-NOT gate CX is defined by

$$CX\,|i\rangle\,|j\rangle = |i\rangle\,|i \oplus j\rangle \qquad i, j = 0, 1.$$

The controlled phase gate CZ is defined by

$$CZ\,|i\rangle\,|j\rangle = (-1)^{ij}\,|i\rangle\,|j\rangle\,.$$

In contrast to the CX gate, CZ is symmetrical in the two input qubits. The Hadamard operator is

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and the Hadamard basis states are

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle).$$

The Bell basis states are

$$\begin{aligned} |B_{00}\rangle &= \tfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = I \otimes I\,|B_{00}\rangle \\ |B_{01}\rangle &= \tfrac{1}{\sqrt{2}}(|01\rangle + |10\rangle) = X \otimes I\,|B_{00}\rangle \\ |B_{10}\rangle &= \tfrac{1}{\sqrt{2}}(|00\rangle - |11\rangle) = Z \otimes I\,|B_{00}\rangle \\ |B_{11}\rangle &= \tfrac{1}{\sqrt{2}}(|01\rangle - |10\rangle) = iY \otimes I\,|B_{00}\rangle\,. \end{aligned}$$

Succinctly we have

$$|B_{cd}\rangle = Z^c X^d \otimes I\,|B_{00}\rangle\,.$$

We will also use the maximally entangled state that combines the Hadamard and standard bases in its Schmidt form:

$$|H\rangle = \frac{1}{\sqrt{2}}\,(|0\rangle\,|+\rangle + |1\rangle\,|-\rangle) = \frac{1}{\sqrt{2}}\,(|+\rangle\,|0\rangle + |-\rangle\,|1\rangle) = CZ\,|+\rangle\,|+\rangle\,.$$

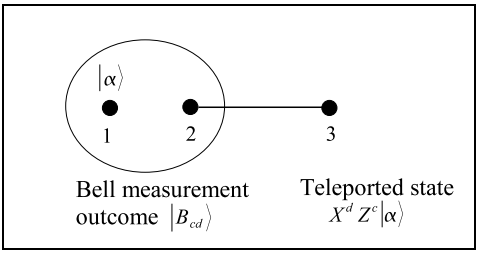


Figure 1. Standard quantum teleportation. The labels 1,2,3 represent three qubits. 1 is in state $|\alpha\rangle$ and 23 are in state $|B_{00}\rangle$. A Bell measurement is performed on 12. If the outcome is $|B_{cd}\rangle$ then qubit 3 acquires the state $X^d Z^c |\alpha\rangle$.

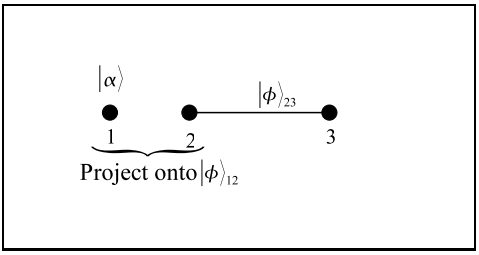


Figure 2. Projecting the three qubit state $|\alpha\rangle_1 |\phi\rangle_{23}$ onto the maximally entangled state $|\phi\rangle_{12}$ results in a single qubit state at 3 given by $\frac{1}{d}|\alpha\rangle_3$ (see lemma 1).

2. Teleportation Based Quantum Computing

Recall standard teleportation [1] as depicted in figure 1 and explained in the caption.

Now introduce the "rotated Bell basis" denoted $\mathcal{B}(U) = \{|B(U)_{cd}\rangle\}$ where for any 1-qubit gate U:

$$|B(U)_{cd}\rangle = U^\dagger \otimes I\, |B_{cd}\rangle\,.$$

A simple calculation shows that if we perform such a rotated Bell measurement (instead of the standard one in figure 1) then the teleported state at 3 is $X^d Z^c U\,|\psi\rangle$ i.e. the gate U has been applied to $|\psi\rangle$ via this teleportation [2]. A particularly neat and general way of performing this calculation is the following. For any dimension d consider any maximally entangled state, written in its Schmidt form as

$$|\phi\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle\,|i\rangle\,.$$

Consider the mathematical projection depicted in figure 2.

Lemma 1: The projection of $|\alpha\rangle_1 |\phi\rangle_{23}$ onto $|\phi\rangle_{12}$ results in the state $\frac{1}{d}|\alpha\rangle_3$ at qubit 3.

Proof: Let $|\alpha\rangle = \sum a_j\,|j\rangle$. Then the projection is

$$\frac{1}{d}(\sum_i \langle i|\,\langle i|)(\sum_{jk} a_j\,|j\rangle\,|k\rangle\,|k\rangle) = \frac{1}{d}\sum_{ijk} a_j \delta_{ij}\delta_{ik}\,|k\rangle = \frac{1}{d}\sum_k a_k\,|k\rangle\,. \quad \square$$

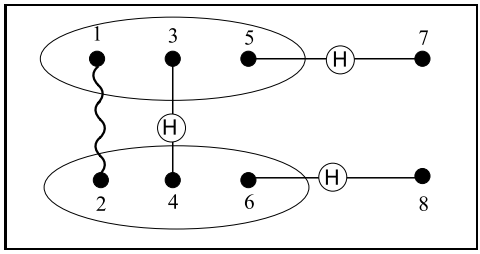


Figure 3. Application of CZ by teleportation with 8 dimensional measurements. The wiggly line connecting 12 denotes an input 2-qubit state $|\psi\rangle$. The lines labelled H denote the maximally entangled state $|H\rangle$. Introduce the 3-qubit Bell measurement corresponding to the basis $\{|000\rangle \pm |111\rangle, |001\rangle \pm |110\rangle, |010\rangle \pm |101\rangle, |100\rangle \pm |011\rangle\}$. If this measurement is performed on qubits 135 and 246 then it can be verified that qubits 78 acquire the state $(P_i \otimes P_j)(H \otimes H)CZ\,|\psi\rangle$ where the Pauli operators P_i and P_j depend on the measurement outcomes.

Next introduce the "rotated $|\phi\rangle$ state": for any unitary operator U in d dimensions define

$$|\phi(U)\rangle = U^\dagger \otimes I\,|\phi\rangle\,. \tag{1}$$

Lemma 2: The projection of $|\alpha\rangle_1 |\phi\rangle_{23}$ onto $|\phi(U)\rangle_{12}$ results in the state $\frac{1}{d}U\,|\alpha\rangle_3$ at qubit 3.

Proof: For any states $|a\rangle$, $|b\rangle$ if we write $\left|U^\dagger a\right\rangle = U^\dagger\,|a\rangle$ then $\left\langle U^\dagger a|b\right\rangle = \langle a|Ub\rangle$. Hence lemma 1 with $|\phi\rangle_{12}$ replaced by $|\phi(U)\rangle_{12}$ immediately gives the result.□

Now the basic idea is to regard the 12-projections in lemmas 1 and 2 as being outcomes of a projective measurement applied to qubits 1 and 2 i.e. we wish to choose a set of d^2 unitaries U_i such that $\{|\phi(U_i)\rangle\}$ form an orthonormal basis of the 12 space. For qubits ($d = 2$) the set $\{I, X, Z, XZ\}$ gives standard teleportation and the set $\{IU, XU, ZU, XZU\}$ gives the rotated Bell basis, reproducing our previous procedure of "gate teleportation" to construct $X^d Z^c U\,|\psi\rangle$ at qubit 3.

Note that lemmas 1 and 2 apply in general dimension d so we can apply 2-qubit gates such as CZ via measurements by teleportation in dimension $d = 4$. For example using $|\phi\rangle = |B_{00}\rangle\,|B_{00}\rangle$ and the 16 operators $U_{ij} = (P_i \otimes P_j)CZ$ where P_i and P_j range over the four standard Pauli operators, we can check that $\{U_{ij}^\dagger \otimes I\,|\phi\rangle\}$ is an orthonormal set and the output teleported state is $(P_i \otimes P_j)CZ\,|\psi\rangle$ for any input 2-qubit state $|\psi\rangle$.

This method of applying CZ requires a 16 dimensional Bell measurement but with more subtle means we can achieve the result with smaller measurements. An illustrative example (taken from [18]) is shown in figure 3.

Remark: Returning to $|\phi(U)\rangle$ in eq. (1), a set of d^2 operators $\{U_i\}$ will have the corresponding set $\{|\phi(U_i)\rangle\}$ being orthonormal iff $\mathrm{Tr}\,(U_i U_j^\dagger) = \delta_{ij}$ i.e. $\{U_i\}$ is a so-called unitary operator basis. For general dimension d many such sets exist, each corresponding to a teleportation scheme. Werner [8] has described a method for constructing a large number of explicit inequivalent examples. However we can go even further and choose a set of $n \geq d^2$ operators U_i such that the set $\{k_i\,|\phi(U_i)\rangle\,\langle\phi(U_i)|\}$ (for some chosen constants $k_i > 0$) form the elements of a (rank 1) positive operator valued measure (POVM) i.e. we ask that

$$\sum_{i=1}^{n} k_i \left|\phi(U_i)\right\rangle \left\langle\phi(U_i)\right| = I_d \otimes I_d.$$

The case of $n = d^2$ (with all k_i's then necessarily equal) reproduces projective measurements as above but for $n > d^2$ we obtain fully valid teleportation schemes in which the Bell measurement is replaced by a POVM. Even in the qubit case $d = 2$ many examples exist and in fact we can have unboundedly large n. (We will not digress here to discuss explicit examples of such constructions.)□

Thus in summary so far, if we have a pool of maximally entangled states we can apply any unitary gate U to any (multi-qubit) input state $|\psi\rangle$ by measurements alone. A significant annoyance (and see more about this later) is that we do not get the exact desired result $U|\psi\rangle$ but instead get $PU|\psi\rangle$ where P is some Pauli operation (on each qubit) depending on the measurement outcome. This is the residue of the randomness of quantum measurement outcomes in our computational formalism.

To perform universal computation it suffices to be able to apply gates from any convenient universal set. We introduce notation for further gates. x and z rotations are defined by

$$R_x(\theta) = e^{-i\theta X} \qquad R_z(\theta) = e^{-i\theta Z}.$$

The following gate will be significant for 1WQC later:

$$W(\theta) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & e^{i\theta} \\ 1 & -e^{i\theta} \end{pmatrix} = HP(\theta) \quad \text{where} \quad P(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}. \tag{2}$$

Any 1-qubit gate U (up to an overall phase) can be decomposed as (Euler angles):

$$U = R_x(\zeta)R_z(\eta)R_x(\xi) \qquad \text{for some } \xi, \eta, \zeta$$

and also [18] as

$$U = W(0)W(\theta_1)W(\theta_2)W(\theta_3) \qquad \text{for some } \theta_1, \theta_2, \theta_3.$$

It is known that CZ together with all 1-qubit operations is a universal set so the sets $\{CZ, R_x(\theta), R_z(\phi) \text{ all } \theta, \phi\}$ and $\{CZ, W(\theta) \text{ all } \theta\}$ are also universal.

3. Adaptive Measurements

We wish to perform arbitrary sequences of gates from a universal set by measurements. For simplicity consider 1-qubit gates: to perform $\ldots U_3U_2U_1|\psi\rangle$ we successively teleport the three gates but instead of the desired result we get $\ldots P_3U_3P_2U_2P_1U_1|\psi\rangle$ where $P_1, P_2, P_3 \ldots$ are Pauli operators depending on the measurement outcomes. To deal with this awkwardness we introduce a new feature: *adaptive choices* of measurements. Let us assume that all 1-qubit operations are x or z rotations. The following Pauli "propagation" relations are easily verified:

$$\begin{aligned} R_x(\theta)X &= XR_x(\theta) \qquad & R_x(\theta)Z &= ZR_x(-\theta) \\ R_z(\theta)X &= XR_z(-\theta) \qquad & R_z(\theta)Z &= ZR_z(\theta). \end{aligned}$$

(Alternatively we could use only $W(\theta)$'s with relations $W(\theta)X = ZW(-\theta)$ and $W(\theta)Z = XW(\theta)$.) Suppose we want to do $\ldots R_z(\beta)R_x(\alpha)\,|\psi\rangle$. The first Bell measurement with outcomes a, b gives $Z^aX^bR_x(\alpha)\,|\psi\rangle$. Now commuting Z^aX^b to the left through $R_z(\beta)$ causes β to change to $(-1)^b\beta$. So if instead of $R_z(\beta)$ we next applied the measurement for $R_z((-1)^b\beta)$ instead, then

$$R_z((-1)^b\beta)Z^aX^bR_x(\alpha)\,|\psi\rangle = Z^aX^bR_z(\beta)R_x(\alpha)\,|\psi\rangle$$

as wanted. This second measurement with basis adapted to measurement outcome b has an outcome c, d say, and the state is now $Z^{a+c}X^{b+d}R_z(\beta)R_x(\alpha)\,|\psi\rangle$. Continuing in this way, adapting measurement bases to earlier measurement results, we get $\ldots X_2^{m_2}Z_2^{n_2}X_1^{m_1}Z_1^{n_1}$(the correct wanted U)$\,|\psi\rangle$. Here $|\psi\rangle$ is generally a state of many qubits and X_i, Z_i are Pauli operations on the i^{th} qubit. The indices m_i, n_i are accumulations (actually bit sums mod 2) of measurement outcomes.

The same idea applies to the 2-qubit gate CZ too where the situation is even better. We have the propagation relations:

$$CZ(Z \otimes I) = (Z \otimes I)CZ \qquad CZ(X \otimes I) = (X \otimes Z)CZ$$

(and similarly for X and Z acting on the second qubit on the LHS's, as CZ is symmetrical). Thus we can propagate Pauli operators through CZ while keeping CZ the *same* i.e. no basis adaption is required!

Remark: More generally for any 1-qubit U we can achieve Pauli propagation by

$$UX = X(XUX) \qquad UZ = Z(ZUZ)$$

where we adapt U to change into XUX or ZUZ. But the latter operators look very different from U and we have (unnecessarily) also preserved the actual identity of the Pauli operations in this propagation. The actual relations we used above for CZ, $R_x(\theta)$, $R_z(\theta)$ and $W(\theta)$ exploit the opportunity of allowing the Pauli operations to change while keeping the adapted operation similar to the original one (e.g. differing only by a sign in the angle). □

If U is the total unitary effect of a gate array on a multi-qubit state $|\psi\rangle$ then using the above methods we are able to generate a state of the form $\ldots X_2^{m_2}Z_2^{n_2}X_1^{m_1}Z_1^{n_1}U\,|\psi\rangle$. In a quantum computation we finally measure (some qubits of) $U\,|\psi\rangle$ in Z_i-bases $\{|0\rangle, |1\rangle\}$ and the presence of the Pauli operations $X_i^{m_i}Z_i^{n_i}$ cause no problems: $Z_i^{n_i}$ has no effect on the measurement outcomes and $X_i^{m_i}$ requires only a simple reinterpretation of the output results – for a single qubit, if $U\,|\psi\rangle = a\,|0\rangle + b\,|1\rangle$ then $XU\,|\psi\rangle = a\,|1\rangle + b\,|0\rangle$. Hence we simply need to reinterpret measurement outcome k_i (0 or 1) as $k_i \oplus m_i$ where m_i is the corresponding X Pauli exponent.

We note a rather curious feature here: these "final" Z measurements are *never adaptive* (being fixed as Z measurements) so they can always be performed *first* before any of the other measurements have been implemented! i.e. the output of the computation can be measured before any of the computation itself has been conducted and the subsequent measurement outcomes simply serve to alter the interpretation of those Z-measurement outcomes!

4. Parallelisable Computations: Clifford Operations

In our measurement based models, distinct measurements always apply to *disjoint* sets of qubits. Hence they all commute as quantum operations and if the measurement basis choices were not adaptive then we could do all the measurements simultaneously, in parallel. The necessity of adaptive choices arose from the Pauli propagation relations, but some operations are special in this regard – the so-called Clifford operations on n-qubits.

Let us introduce the Pauli group $\mathcal{P}_n$ on n qubits, defined as the group generated (under multiplication) by n-fold tensor products of $\pm I$, $\pm iI$, X and Z. For example $\mathcal{P}_3$ has elements such as $X \otimes Z \otimes Z$, $-iZ \otimes Y \otimes I$, $I \otimes I \otimes X$ etc. An n-qubit unitary operation C is defined to be a Clifford operation if $C\mathcal{P}_n C^\dagger = \mathcal{P}_n$ i.e. $C\mathcal{P}_n = \mathcal{P}_n C$ i.e. for every Pauli operation $P \in \mathcal{P}_n$ there is another $P' \in \mathcal{P}_n$ such that $CP = P'C$. Hence for any Clifford operation C we can propagate Pauli operations across C while C stays the same i.e. no adaption is needed (but the Pauli operation generally changes). For each n the Clifford operations evidently form a group, called the Clifford group.

We have already noted that CZ is a 2-qubit Clifford operation. Also the Hadamard operation has the Clifford property since $HX = ZH$ and $HZ = XH$. (For any prospective C we need only check the propagation of X and Z at each qubit to verify the full Clifford property.) Indeed we can give an explicit description of the Clifford group on n qubits [9]. Introduce the $\pi/4$ phase gate:

$$P_{\pi/4} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}.$$

Then we have:

Theorem:[9] The Clifford group on n qubits is generated by Z, H, $P_{\pi/4}$ and CX acting in all combinations on any of the qubits (i.e. arbitrary arrays of these gates). □

Hence any array of Clifford gates $C_k \ldots C_2 C_1 |\psi\rangle$ ($|\psi\rangle$ of n qubits) may be implemented in TQC in one parallel layer of measurements. We get a result of the form $P_k C_k \ldots P_2 C_2 P_1 C_1 |\psi\rangle$ where P_i are all Pauli operations on the qubits. Commuting them all out we get $X_n^{a_n} Z_n^{b_n} \ldots X_1^{a_1} Z_1^{b_1} (C_k \ldots C_1) |\psi\rangle$ where the indices a_i, b_i depend on the measurement outcomes and the Clifford propagation relations. The collection of maximally entangled states used in all the teleportations can also be manufactured in parallel (e.g. apply CX's to many pairs $|0\rangle |0\rangle$) so the entire quantum process requires only a *constant* amount of quantum parallel time for any n, in contrast to the corresponding gate array whose depth generally increases with n.

However there is a further subtle point here: in addition to the constant parallel time of the quantum process, the computation of the Pauli exponents a_i, b_i requires a further *classical* computation, which is actually the bitwise sum of selected measurement outcomes. Hence this classical computation can be done as a parallel computation of *log* depth: to sum k bits $i_1 \ldots i_k$ we first sum all pairs $i_1 \oplus i_2$, $i_3 \oplus i_4$, ... in parallel and then pairs of the results etc. At each stage the number of bits is halved so we require $\log k$ layers to reach the final result. Returning to the gate array of Clifford operations, it is known that any Clifford

operation on n qubits can be represented as an array of $O(\log n)$ depth i.e. the same as the total quantum plus classical depth of the full measurement based implementation. But the virtue of the measurement based approach is to take a fully quantum process (in this case, the Clifford gate array of log depth) and recast it as a quantum process of constant depth plus a classical computation (of log depth in our case) i.e. we separate the original process into a quantum and classical part while suitably "minimising" the quantum part. This kind of restructuring is significant in a scenario where quantum and classical computation are regarded as separate (or even incomparable) resources (c.f. further discussion in section 8 below) and we can ask interesting questions such as: what is the "least" amount of quantum "assistance" needed to supplement classical polynomial time computation in order to capture the full power of quantum polynomial time computations?

Remark: It is known that arrays of quantum Clifford operations can be classically efficiently simulated. This is the Knill-Gottesman (KG) theorem [10] which asserts the following: consider any array of Clifford gates on n qubits, each initialised in state $|0\rangle$. Let $\mathcal{P}$ be the probability distribution resulting from measuring (some of) the output qubits in the Z basis. Then there is a classical (probabilistic) computational process which runs in time polynomial in n, which also has $\mathcal{P}$ as its output distribution. Furthermore according to [11] this classical simulation of any Clifford array's output can always be performed in log depth. In view of this, one may question the significance of our result above, that any such Clifford array can be reproduced as a classical (log depth) process *plus* a quantum process of constant depth. However there is an essential difference in the two representations of the Clifford array: the KG simulation results in a *purely classical* output (sample of $\mathcal{P}$) whereas our classical-quantum separated simulation results in a *quantum* state as output i.e. we are simulating the quantum process itself rather than just the classical output of some measurement results. Indeed our result makes an interesting statement about *quantum* properties of Clifford arrays, viz. that the essential quantum content can be "compressed down" into constant quantum depth, which is not provided by the KG result. □

5. The "One-Way" Quantum Computer

We now move on to describe our second measurement based computational model – the 1WQC of Raussendorf and Briegel [5,6] At first sight it looks rather different from TQC but we will see later that the models are in fact very closely related.

Consider a rectangular (2 dimensional) grid of $|+\rangle$ states as shown in figure 4 . We apply CZ to each nearest neighbour pair (in horizontal and vertical directions). These CZ's all commute so for any grid size they can all be applied in parallel, as a process of constant quantum depth. The resulting state is an entangled state of many qubits, called a *cluster state*. We will also use one dimensional cluster states constructed in the same way, but starting from a one dimensional array of $|+\rangle$ states.

The 1WQC is based on the following facts that will be elaborated below. Any quantum gate array can be implemented as a pattern of 1-qubit measurements

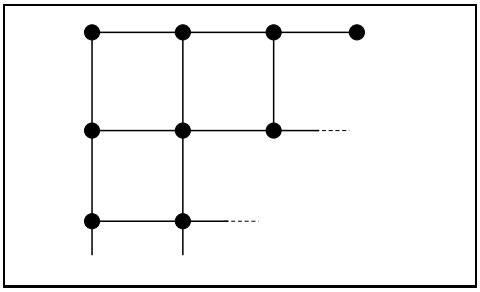


Figure 4. Construction of the cluster state. The dots represent a grid of $|+\rangle$ states and a line connecting a pair of dots represents the application of CZ to the corresponding pair of qubits.

on a (suitably large two dimensional) cluster state. The only measurements used are in the bases $M_z = \{|0\rangle, |1\rangle\}$ and $M(\theta) = \{|0\rangle \pm e^{i\theta}|1\rangle\}$ for some θ. $\theta = 0$ corresponds to an X basis measurement. Measurement outcomes are always labelled 0 or 1 and they are always uniformly random. As in teleportation, quantum gates are implemented only up to Pauli corrections $X^a Z^b$ where a and b depend on the measurement outcomes. (These Pauli corrections are called bi-product operators in the 1WQC literature). Hence we'll get the same feature of adaptive measurements that we saw in TQC. The name "one-way" quantum computer arises from the feature that the initial resource of the pure cluster state is irreversibly degraded as the computation proceeds in its layers of measurements.

To illustrate these ingredients we give some explicit examples of measurement patterns for 1-qubit gates (where 1-dimensional cluster states suffice). Our first example is taken from [6]. We noted previously that any 1-qubit U can be expressed as

$$U = R_x(\zeta)R_z(\eta)R_x(\xi) \quad \text{for some } \xi, \eta, \zeta.$$

To apply U to $|\psi\rangle$ by the 1WQC method we start with $|\psi\rangle$ in a line with four $|+\rangle$ states, as shown in figure 5. (Later we will see how to eliminate explicit use of $|\psi\rangle$ here, starting with only $|+\rangle$ states). Entangle all neighbouring pairs with CZ and subsequently measure qubits 1,2,3 and 4 adaptively in the bases shown in figure 5. Then it may be shown that qubit 5 acquires the state $X^{s_2+s_4}Z^{s_1+s_3}U|\psi\rangle$, where s_i is the outcome of the measurement on qubit i. Note that the measurement bases are adaptive and in the above pattern the measurements must be performed in numerical sequence.

As a second example consider the pattern of figure 6. Then qubit 3 acquires the state $Z^{s_1}X^{s_2}|\psi\rangle$ so we have a process which is very similar to teleportation from 1 to 3. The two X measurements may be done in parallel but we can also consider this process as a sequence of two identical steps: given a state $|\psi\rangle$, adjoin $|+\rangle$, entangle with CZ and then X-measure the first qubit, giving an outcome s. Qubit 2 is then left in state $X^s H|\psi\rangle$. Thus the two-step chain of figure 6 can be analysed as $(X^{s_2}H)(X^{s_1}H)|\psi\rangle = X^{s_2}Z^{s_1}HH|\psi\rangle = X^{s_2}Z^{s_1}|\psi\rangle$ (where we have used the Pauli propagation relations for H).

Figure 7 shows a single step operation for the general measurement basis $M(\theta)$. Qubit 2 is then left in state $X^{s_1}W(-\theta)|\psi\rangle$ (with $W(\theta)$ as defined in eq. (2)). Indeed if we are not concerned with issues of parallelisability then any 1-

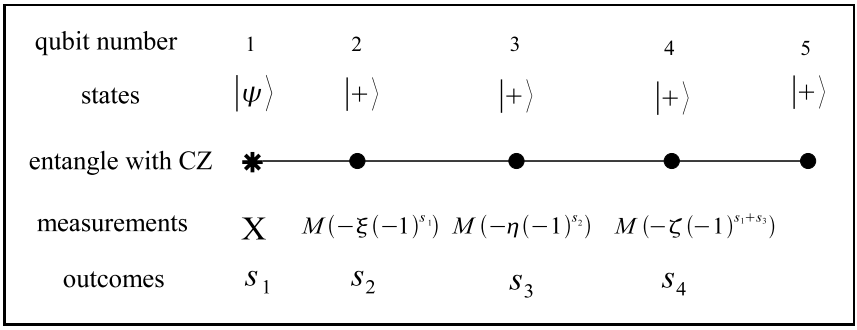


Figure 5. 1WQC implementation of a 1-qubit unitary U on $|\psi\rangle$. ξ, η, and ζ are Euler angles for U. The leftmost qubit, denoted by a star, is set in state $|\psi\rangle$ and extended by a row of four $|+\rangle$ states denoted by dots. CZ operations are then applied, denoted by connecting lines. Next, measurements are applied in the designated bases with outcomes s_i. Hence the measurements must be carried out adaptively from left to right. As a result of this process the rightmost (unmeasured) qubit is left in state $X^{s_2+s_4}Z^{s_1+s_3}U|\psi\rangle$.

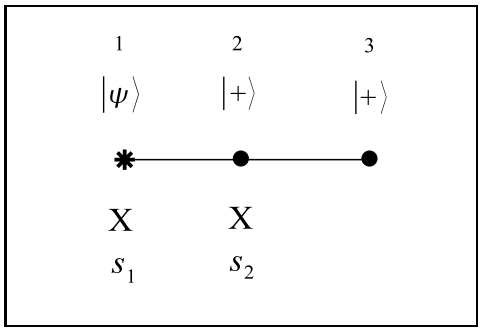


Figure 6. See caption of figure 5 for explanation of pictorial notations. After the two X measurements, qubit 3 is left in state $Z^{s_1}X^{s_2}|\psi\rangle$ where $|\psi\rangle$ was the input state at qubit 1.

dimensional measurement pattern may be viewed as a sequential application of the process of figure 7 applied repeatedly. This shows that the unitary operation $W(\theta)$ plays a fundamental role in 1WQC. For example in figure 5 we can reorder all operations as follows: (entangle 12, measure 1), (entangle 23, measure 2), etc. This is because the 12-entangling operation and first measurement both commute with all subsequent entangling operations and measurements.

In a similar way it is now straightforward to see how gates may be concatenated. Suppose we wish to apply U_1 and then U_2 to an input qubit $|\psi\rangle$. Measurement pattern 1 for U_1 has an output qubit (e.g. qubit 5 in figure 5) which is the input qubit for the measurement pattern of U_2. But all measurements in pattern 1 commute with all entangling operations of pattern 2. Hence we can apply all entangling operations (for both patterns) first, to get a longer single cluster state and then apply the measurements. Furthermore some measurements in pattern 2 could even be performed *before* those in pattern 1 if their basis choice does not depend on pattern 1 outcomes.

Finally we can eliminate the input state $|\psi\rangle$ from the above descriptions, to get a formalism based entirely on a starting state that's a fully standard cluster

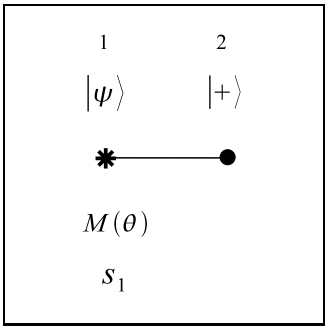


Figure 7. Effect of a single $M(\theta)$ measurement in 1WQC. Qubit 2 is left in state $X^{s_1}W(-\theta)\,|\psi\rangle$.

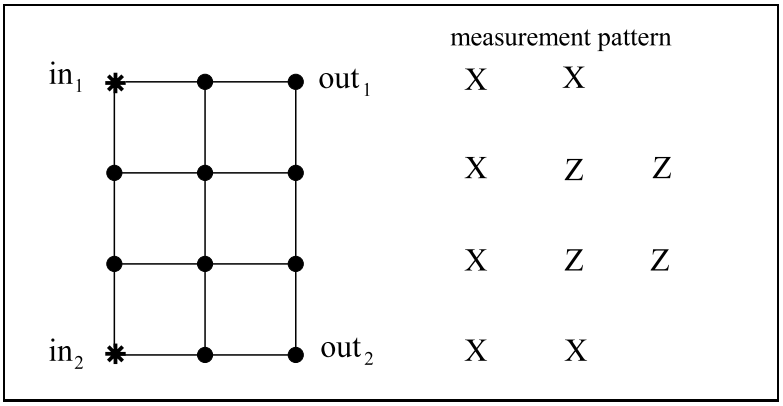


Figure 8. Application of CZ gate in 1WQC. The 2-qubit input state $|\psi_{\text{in}}\rangle$ is placed at sites labelled in_1 and in_2. Dots denote $|+\rangle$ states and connecting lines denote application of CZ for cluster state generation. If the measurement pattern shown at the right is applied at the sites, then only sites out_1 and out_2 remain unmeasured and contain $(P_1 \otimes P_2)CZ\,|\psi_{\text{in}}\rangle$ where $P_1 \otimes P_2$ is a Pauli operation that depends on the measurement outcomes.

state (of slightly longer length): since we can implement any 1-qubit U we can take our input starting state to be $|+\rangle$ and prefix the desired process with an initial measurement pattern for a unitary operation that takes $|+\rangle$ to $|\psi\rangle$.

Above we have considered only 1-qubit gates but the formalism may be generalised (using 2-dimensional cluster grids) to incorporate 2-qubit gates. For universal computation it suffices to be able to implement just the CZ gate (in addition to 1-qubit gates). An explicit measurement pattern for CZ is shown in figure 8. It should also be noted that measurement patterns are not unique and subject to various approaches for their invention.

5.1. Role of Z Measurements

The CZ gate requires use of a 2-dimensional grid whereas 1-qubit gates require only 1-dimensional clusters. Hence measurement patterns for general gate arrays, implemented on a suitably large 2-dimensional cluster, will generally have some extraneous sites not used in the measurement patterns. Z-measurements are used to delete such extraneous sites.

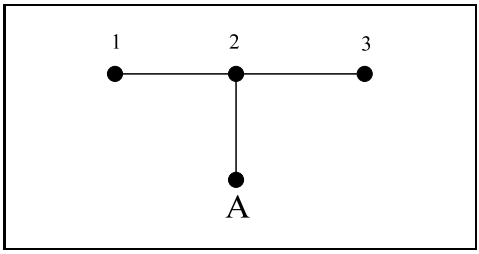


Figure 9.

To illustrate the principle consider a cluster state with irregular shape in figure 9. Suppose we wish to use only the linearly arranged sites 1,2,3 i.e. we want to delete site A. Consider Z_A, a Z measurement at site A, with outcome k_A. To see its effect recall that the pattern in figure 9 is obtained by applying CZ operations to a set of four $|+\rangle$ states, located at the sites. But measurement Z_A commutes with CZ_{12} and CZ_{23} so starting with the four $|+\rangle$ states we can first do CZ_{A2} and then Z_A before CZ_{12} and CZ_{23}. But $CZ_{A2}|+\rangle|+\rangle = |0\rangle|+\rangle + |1\rangle|-\rangle$ so after the Z_A measurement (with outcome k_A) the sites 1,2,3 contain $|+\rangle\left|(-1)^{k_A}\right\rangle|+\rangle$ i.e. the effect of the Z_A measurement is to have used the Hadamard basis state $\left|(-1)^{k_A}\right\rangle$ at site 2 instead of the standard $|+\rangle$ (with no site at A) and then entangle as usual. Next consider the basic 1WQC single step of figure 7 with $|-\rangle$ instead of $|+\rangle$ at site 2: we start with $|\psi\rangle|-\rangle$, apply CZ, then do $M(\theta)$ at qubit 1 for output at qubit 2. But this is the same as the following: start with $|\psi\rangle|+\rangle$, then (i) apply CZ, (ii) do $M(\theta)$ on qubit 1, then (iii) perform the unitary operation $I_1 \otimes Z_2$. This is identical to the previous process because (iii) commutes with (i) and (ii) so we can change the order to (iii), then (i) then (ii) and note that $Z_2|+\rangle = |-\rangle$. A similar argument applies to each neighbour of A if there is more than one. In summary, we see that a Z_A measurement (outcome k_A) at a site A adds in an extra Z^{k_A} Pauli correction at all neighbouring sites of A, in addition to the usual Pauli operations arising from measurement patterns in a standard cluster state that had site A absent from the start.

Z measurements also have a second role: as in TQC they are applied to a final state of a computation to produce the classical output results. So just as in TQC we have the curious feature that these "final" Z measurements are always non-adaptive and can be applied first, before any of the computation itself has been implemented!

To briefly summarise the 1WQC model, we have seen that any quantum gate array can be translated into a pattern of 1-qubit measurements on a suitably large 2-dimensional cluster state. Choices of measurement bases are generally adaptive i.e. possibly depending on previous measurement outcomes. Thus the 1-qubit measurements are organised into layers and the measurements within each layer can be done simultaneously in parallel. The output qubits (always measured in the Z basis) can always be done in the first layer. This temporal sequencing "cuts across" the temporal sequence of the original gate array – all gates are generally done "partially" in each layer and simultaneously built up as the layers accumulate. Just as in TQC, arrays of Clifford operations can always be fully implemented with only *one* layer of measurements.

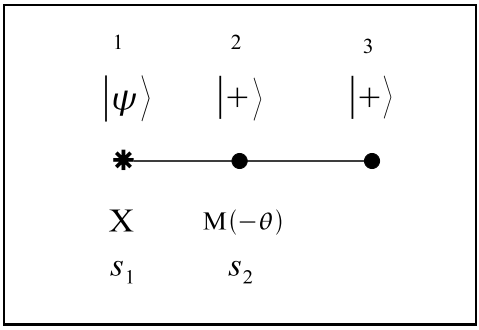


Figure 10. The process above leaves qubit 3 in state $X^{s_2} Z^{s_1} R_x((-1)^{s_1}\theta)\,|\psi\rangle$. Thus to implement a given x rotation the sign of the angle θ is adaptive, depending on the measurement outcome s_1.

6. Further Features of 1WQC

6.1. A Further Parallelisability Result

Any polynomial sized quantum gate array can be implemented in 1WQC using at most a polynomial number of measurement layers (c.f. section 8 later). Also we have seen that any array of Clifford operations can be implemented with just one layer of measurements. This suggests the following interesting question: which classes of quantum gate arrays can be implemented in 1WQC with constraints on the number of measurement layers e.g. using 2 or 3 or a logarithmic number of layers? Does the latter include all polynomial time quantum computation? Although very little is known about such questions, we have the following result of Raussendorf and Briegel [6,12]. Its proof depends on more subtle properties of Pauli propagation relations for particular operations.

Theorem: Any gate array using gates from the set $\{CX, R_x(\theta) \text{ all } \theta\}$ or the set $\{CX, R_z(\theta) \text{ all } \theta\}$ can be implemented with just two measurement layers.

Remark: Neither of these sets is believed to be universal although it is known that CX with all y-rotations *is* universal [13].

Proof of theorem: We give a proof for x-rotations. (The case of z-rotations is similar). We use the following three (easily verified) facts: (i) any single $M(\theta)$ measurement (e.g. as shown in figure 7) generates only I or X Pauli corrections. They may become Z's only after propagation through other gates. (ii) when Pauli operations are commuted across the Clifford operation CX, X propagates only to X's (and Z propagates only to Z's) although the Pauli operator may spread from one qubit onto two qubits. (iii) The R_x gate may be implemented in 1WQC using the measurement pattern shown in figure 10.

Now consider any gate array $\ldots G_3 G_2 G_1$ where each G_i is a CX or an x-rotation gate. Set up its corresponding measurement pattern (using figure 10 for each x-rotation gate). CX is a Clifford operation and involves no adaptive measurements. In the first measurement layer we perform all CX pattern measurements and all X measurements of the R_x gates (i.e. all qubit 1 measurements in figure 10). This produces some extraneous Pauli operations and leaves only the $M(-\theta)$ nodes unmeasured (i.e. all qubit 2's in figure 10). Next commute all

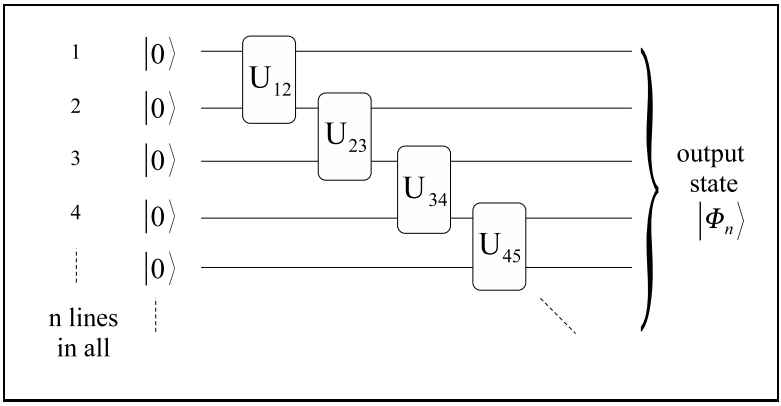


Figure 11. A generalised 1-dimensional cluster state $|\Phi_n\rangle$. Starting with a row of n qubits each in state $|0\rangle$ we sequentially apply $U_{i,i+1}$ to qubits $i, i+1$ as shown in the above gate array having a ladder-like structure.

these Pauli's out to the left hand end of the gate array. This commutation leaves all CX's unchanged (as CX is Clifford) but some x-rotation angles acquire an unwanted minus sign (c.f. propagation relations for x-rotations given previously). Thus reset these altered angle signs so that the $M(-\theta)$ measurements again give the correct designated gates i.e. make the adaptive choice of bases for the next measurement layer. Finally in the second layer, perform all the $M(-\theta)$ measurements. We again get some further extraneous Pauli operations generated. By (i) we get only I or X, but these can be harmlessly commuted out to the left since X commutes with x-rotations and CX preserves X's in its propagation relations (i.e. no Z's are generated). □

6.2. Non-Universality of One-Dimensional Clusters

We have seen that 1WQC with two-dimensional cluster states is universal for quantum computation. Nielsen and Doherty [14] have shown that any 1WQC process based on only 1-dimensional cluster states can be simulated classically efficiently i.e. in polynomial time in the number of qubits. Thus the universality of any such model would imply that quantum computation is no more powerful than classical computation.

The argument may be paraphrased as follows (ignoring technical issues of precision of the simulation). Consider any state of the form shown in figure 11. (The standard cluster state is obtained by choosing each $U_{i,i+1}$ to be $CZ(H \otimes H)$). Consider now any sequence of 1-qubit measurements such that for later measurements, the choice of measurement basis and even the choice of qubit used, may both depend on outcomes of earlier measurements. Then the whole process may be classically simulated in polynomial (in n) time i.e. the resulting probability distribution of outcomes may be sampled by classical means in polynomial time.

The proof runs as follows. At a general stage suppose measurements on lines $a, b, \ldots, p, q$ have already been simulated and the chosen sample outcomes are $a', b', \ldots, p', q'$ respectively. Even though these measurements may have been chosen adaptively, once the outcomes have been specified (as $a', b', \ldots, p', q'$) we can

perform the measurements in any order we wish to compute the joint probability $P(a', b', \ldots, p', q')$ of the designated outcome string. Now suppose the next measurement is on line k. To sample its outcome distribution we need to know

$$P(k'|a', b', \ldots, p', q') = \frac{P(a', b', \ldots, k', \ldots, p', q')}{P(a', b', \ldots, p', q')}.$$

Here $a', b', \ldots, p', q'$ have their fixed values and we take $k' = 0, 1$ separately. Having computed this probability we sample k' to give a definite value, and continue in the same way for the next measurement on some line, l say.

To show that this whole process is efficient, we need only show that $P(a', b', \ldots, p', q')$ can be computed in poly(n) time for any chosen set of values $a', b', \ldots, p', q'$, on any chosen set of lines. Without loss of generality suppose these lines are listed in increasing order of occurrence from line 1 at the top. Note that we can regard a measurement on any line, a say, as occurring immediately after $U_{a,a+1}$ and before $U_{a+1,a+2}$. Let ρ_m denote the reduced state of any line m at the position in between the application of $U_{m-1,m}$ and $U_{m,m+1}$, and let $\sigma_{m,m+1}$ denote the application of $U_{m,m+1}$ to $\rho_m \otimes |0\rangle \langle 0|$.

Starting at the top we compute σ_{12} as U_{12} on $|0\rangle |0\rangle$. Then compute ρ_2 by partial trace, tracing out system 1 from σ_{12}. Then compute σ_{23} as U_{23} on $\rho_2 \otimes |0\rangle \langle 0|$ etc. continuing until $\sigma_{a,a+1}$ where the first measurement occurs. At this point we do not compute ρ_{a+1} as the partial trace above but instead, apply to $\sigma_{a,a+1}$ the projector corresponding to measurement outcome a', obtaining a subnormalised state ρ_{a+1} on line $a+1$. In fact $\mathrm{Tr}\, \rho_{a+1}$ is the probability of getting outcome a' in the a-line measurement. We continue in this way computing the reduced states of the lines successively (using the measurement projector at any measured line and partial trace at any unmeasured line) until we have applied the last (q') measurement projector. The trace of this final resulting state is then $P(a', b', \ldots, p', q')$. At each stage of this calculation we need to hold the state of at most two lines (i.e. not n lines with its exponentially large description) and we pass through the ladder n times, once for each successive measurement. Hence the whole calculation is completed in time polynomial in n. This feature of the calculation is a consequence of the special ladder-like structure of figure 11 and it does not apply to general gate arrays.

7. Relationships Between the TQC and 1WQC Models

The two models have several similarities – both are based on measurements as computational steps and both have the awkward feature of supplementing desired gates with unwanted Pauli operations. But there are also some essential differences: TQC uses (Bell) measurements on 2 or more qubits whereas 1WQC uses only 1-qubit measurements. 1WQC starts with a cluster state having multi-partite entanglement across *all* the qubits whereas TQC is based on a state comprising only *bipartite* entangled pairs.

The models can be related in several different ways. We will discuss three of them. The relationships serve to improve our understanding of 1WQC and its

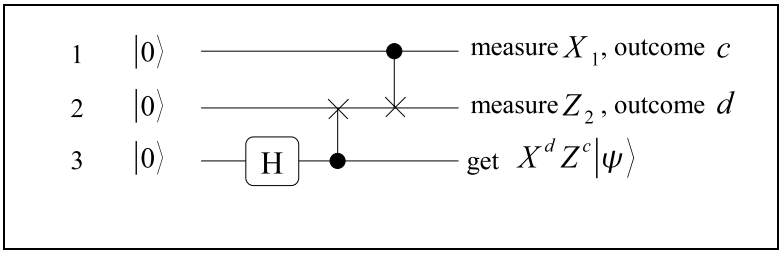


Figure 12. Teleportation as entangling operations and 1-qubit measurements. The diagram is read from left to right. The vertical lines denote CX operations with control and target marked as • and X respectively. Note that the two CX operations, with the shown opposite orientations, commute. Alternatively we could have put H on line 2 and the first CX the other way up, but then the two CX's would not commute.

measurement patterns. For TQC the relation between the measurement and the desired gate is already transparent.

Our first way of relating TQC and 1WQC was proposed by Aliferis and Leung[15]. The basic idea is to identify suitable *pairs* of consecutive 1-qubit measurements in 1WQC with a Bell measurement of a teleportation. This approach is strongly suggested by patterns such as the one in figure 6. With reference to the basic teleportation scheme in figure 1 we note the following. CX transforms Bell states into product states:

$$CX\,|B_{ij}\rangle = \left|(-1)^j\right\rangle |i\rangle$$

(where the first ket is a $|\pm\rangle$ state according to the given sign). Hence the Bell measurement on 12 can be performed by the entangling operation CX followed by the 1-qubit measurements X_1 and Z_2. Similarly the $|B_{00}\rangle$ state of 23 is $CX\,|+\rangle_2\,|0\rangle_3$ which we can alternatively write as $CX\,|+\rangle_3\,|0\rangle_2$ (where subscripts denote the qubit number and we adopt the notation for the asymmetrical CX operation that the first qubit listed is the control qubit). Thus teleportation fully decomposes into entangling operations and 1-qubit measurements as shown in figure 12. So, in view of figure 12 teleportation can be interpreted as: start with $|\psi\rangle_1\,|0\rangle_2\,|+\rangle_3$, entangle suitably with CX's, then do X_1 and Z_2 measurements, which is structurally just like the 1WQC paradigm. But we have a slight mismatch in choice of primitives: TQC uses CX and $|B_{00}\rangle$ states whereas 1WQC is based on CZ and $|H\rangle = CZ\,|+\rangle\,|+\rangle$ states, so the correspondence involves a sprinkling of Hadamard operations to interconvert these ingredients.

Extending this idea, we find that other pairs of of measurements in 1WQC (such as $X_1 M(\theta)_2$ and $M_1(\theta)X_2$) can be interpreted as rotated Bell measurements, but only special pairs of such consecutive 1-qubit measurements can be fused together to form Bell measurements. We refer to [15] for further details that we will not need here. Although we are able to reconstruct rotated Bell measurements for the 1WQC implementation of a full universal set of gates, this interpretation of 1WQC has the drawback that single 1-qubit measurements individually cannot be interpreted in terms of TQC.

Our second relationship between TQC and 1WQC, proposed by Childs et al.[16] and Jorrand et al.[17], is a further development of the ideas in figure 12. As noted in the caption, the two CX operations commute. Also the X_1 measurement

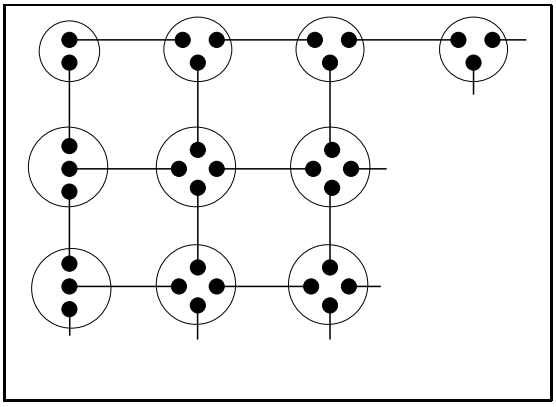


Figure 13. Valence bond solid or matrix product state for the cluster state. Each "valence bond" line denotes the maximally entangled state $|H\rangle$ of two qubits. Each site (circled) has 4 or 3 or 2 qubits. At each site we consider the two dimensional subspace spanned by the two kets of "all zeroes" and "all ones".

commutes with all subsequent operations on 23. Thus we can change the order of actions to: (entangle 12, measure X_1), then (entangle 23, measure Z_2), obtaining a sequence of two operations of the same form viz. (entangle 12, measure 1) to obtain a state at 2. As noted in figure 7, it is exactly this kind of operation that drives 1WQC too, so we can regard it as a common fundamental primitive underlying both models (and sometimes called "one-bit teleportation").

A slight awkwardness in figure 12 is the lack of uniformity of actions: the CX's act in opposite orientations (necessary for commutativity) and we have a single H gate as well. But this can be easily remedied: instead of the usual Bell state based teleportation scheme we consider teleportation with maximally entangled state $|H\rangle = CZ\,|+\rangle\,|+\rangle$ at 23 and its associated Bell measurement on 12 given by the basis

$$\{I \otimes I\,|H\rangle\,, X \otimes I\,|H\rangle\,, Z \otimes I\,|H\rangle\,, XZ \otimes I\,|H\rangle\}. \tag{3}$$

Then note that CZ maps this basis to $\{|+\rangle\,|+\rangle\,, |+\rangle\,|-\rangle\,, |-\rangle\,|-\rangle\,, |-\rangle\,|+\rangle\}$ so the Bell measurement is equivalent to applying CZ and then measuring X_1 and X_2. Also unlike CX, CZ is symmetrical so the picture as in figure 12 for *this* teleportation process is fully uniform. The two "one-bit teleportations" are now identical, in fact corresponding exactly to the process in figure 6, implemented sequentially.

7.1. Matrix Product State Relationship of TQC and 1WQC

Our third and most remarkable connection between TQC and 1WQC, proposed by Verstraete and Cirac[18], is based on the formalism of so-called valence bond solids or matrix product states. In this correspondence each single 1-qubit measurement of 1WQC will be interpreted in terms of a full single teleportation.

Consider a 2-dimensional grid of states $|H\rangle = CZ\,|+\rangle\,|+\rangle$ as shown in figure 13. Let $|\text{grid}\rangle$ denote the total state of all the qubits. At each site in the figure consider the two dimensional subspace spanned by $\{|00\ldots0\rangle\,, |11\ldots1\rangle\}$ and the associated projector, renaming these two basis states as $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$:

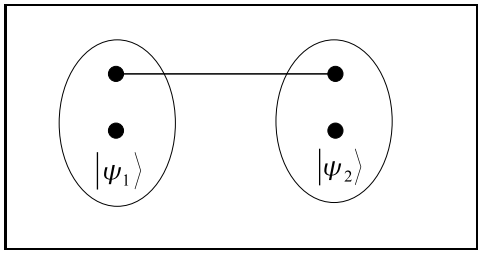


Figure 14. Dots denote qubits and the connecting line denotes the maximally entangled state $|H\rangle$. If we project each circled site to span$\{|00\rangle, |11\rangle\}$ then the resulting state (after normalisation) is $CZ\,|\psi_1\rangle\,|\psi_2\rangle$.

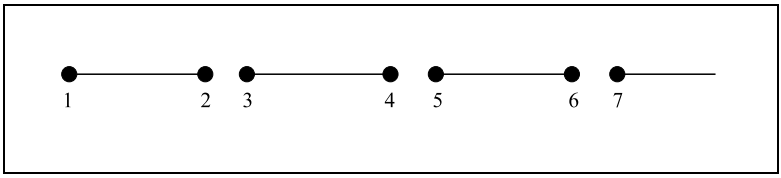


Figure 15. One dimensional valence bond solid.

$$\Pi = \left|\tilde{0}\right\rangle \langle 00\ldots 0| + \left|\tilde{1}\right\rangle \langle 11\ldots 1|\,.$$

Applying Π to $|\text{grid}\rangle$ we obtain a state with a single qubit at each site (and subnormalised because of the projection).

Lemma 3: The multi-qubit state $\Pi\,|\text{grid}\rangle$ (after normalisation) is precisely the 1WQC cluster state.

Proof: We first note the fact of figure 14. Now consider a 1-dimensional grid of $|H\rangle$ states as in figure 15. Apply Π at each node. Each bond is already $CZ\,|+\rangle\,|+\rangle$ by definition. By the fact in figure 14 the projections at sites 23 and 45 simply serve to apply CZ between qubits 2 and 5. Hence the whole projected state is just CZ applied to all connecting pairs in $|+\rangle\,|+\rangle\ldots|+\rangle$ i.e. the 1-dimensional cluster state. This argument easily generalises to the 2-dimensional geometry of figure 13.□

Next consider using $|H\rangle$ states for TQC via application of rotated versions of the associated basic Bell measurement eq. (3). For clarity of the essential idea, consider the 1-dimensional case of figure 15. Let us calculate the rotated Bell basis corresponding to the 1-qubit gate

$$W(-\theta) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & e^{-i\theta} \\ 1 & -e^{-i\theta} \end{pmatrix}.$$

The basis is given by

$$|a\rangle = (W(-\theta)^\dagger \sigma_a) \otimes I\,|H\rangle$$

where $a = 0, 1, 2, 3$ and $\sigma_0 = I$, $\sigma_1 = X$, $\sigma_2 = Z$, $\sigma_3 = XZ$. A direct calculation gives the first two states as

$$|a = 0\rangle = \tfrac{1}{\sqrt{2}}(|00\rangle + e^{i\theta}\,|11\rangle)$$
$$|a = 1\rangle = \tfrac{1}{\sqrt{2}}(|00\rangle - e^{i\theta}\,|11\rangle)$$

and the remaining two span the orthogonal complement of span$\{|00\rangle, |11\rangle\}$ at the site. Thus remarkably these first two Bell states lie within the 1-qubit subspace determined by the Π-projection. Furthermore this part of the Bell measurement corresponds precisely to the measurement basis $M(\theta) = \{|\tilde{0}\rangle \pm e^{i\theta} |\tilde{1}\rangle\}$ on the projected site i.e. on the cluster state. (Note that for other more general 1-qubit gates U, the corresponding rotated Bell basis states do not generally lie in a simple way relative to the Π-projected subspaces.)

Stated otherwise, $M(\theta)$ measurements on the cluster state (which is the basic ingredient of 1WQC, c.f. figure 7) can be thought of as teleportations in the TQC formalism, where the teleportations always produce one of the $a = 0, 1$ outcomes (and not $a = 2, 3$) i.e. teleportations that have been "cut down" by the Π projection. In a similar way all other 1WQC ingredients (viz. Z measurements and the CZ measurement pattern) can be seen as descendants under the Π projection of teleportations on a valence bond grid state $|\text{grid}\rangle$. We omit further details which may be found in [18].

8. Measurement Based Models and Computational Complexity

The gate array model of quantum computation provides a transparent formalism for the theoretical study of quantum computation and its computational complexity features compared to classical computation. So why should we bother with further exotic models such as the measurement based models? Indeed our measurement based models are readily seen to be polynomial time equivalent to the gate array model i.e. each model can simulate the other with only a polynomial (i.e. modest) overhead of resources (number of qubits and computational steps). To see this first recall that the standard gate array model (allowing measurements only at the end and only in the Z basis) can be easily generalised to allow measurements along the way with subsequent choices of further gates and measurements being allowed to depend on earlier measurement outcomes. Indeed consider a measurement in a basis $\{U|0\rangle, U|1\rangle\}$ on a qubit B applied during the course of a gate array process. To regain the standard gate array paradigm, for each such measurement we adjoin an extra ancillary qubit A, initially in state $|0\rangle$ and replace the measurement by the following: apply $U^\dagger$ to B and the apply CX to qubits BA. This simulates a coherent representation of the measurement in which qubit A plays the role of a pointer system. Subsequent gates that depend on the measurement outcome are replaced by a corresponding controlled operation, controlled by the state of A (written in the Z basis). In this way we purge all intermediate measurements from the body of the array and a measurement of each ancilla in the Z basis at the end results in a standard gate array process which is equivalent to the given non-standard one.

Using the above technique any 1WQC process is easily converted into an equivalent (standard) gate array process. We first build the required cluster state using an array of CZ gates acting on $|+\rangle = H|0\rangle$ states and introduce an ancilla A for each 1-qubit 1WQC measurement. For each $M(\theta)$ measurement we introduce an extra gate $W^\dagger(\theta)$ which transforms the $M(\theta)$ basis to the standard basis. The overhead in number of qubits and gates in this simulation is at most linear.

Conversely given any gate array (based say on one of our previously considered universal sets of gates) we have seen how it can be translated into a measurement pattern on a suitably large cluster state. If K is the largest size of the 1WQC measurement pattern for any gate in our universal set then the number of qubits and computational steps increases by at most a factor of K i.e. the resource overhead is again linear.

Polynomial time equivalence of computational models is important in computational complexity theory because such models have the same class of polynomial time computations. But polynomial time equivalence does not preserve more subtle structural features of computations, such as parallelisability. Indeed already in the context of classical computation it is well known that the (one tape) turing machine model is polynomial time equivalent to the (classical) gate array model yet the turing machine model does not even have a natural notion of parallalisability at all, whereas the gate array model does! (i.e. doing gates simultaneously in parallel).

In contrast to the quantum gate array model, the formalism of measurement based models offers new perspectives for parallelisability issues. We have already noted the fundamental feature that measurements on different subsystems of an entangled state always commute so long as the choice of measurement is not adaptive i.e. not dependent on the outcome of another measurement. We have seen examples of processes which are inherently sequential for gate arrays (e.g. sequences of Clifford gates) that become parallelisable in the measurement based models.

The measurement based models have a further novel feature: they provide a natural formalism for separating a quantum algorithm into "classical parts and quantum parts". In contrast, in the gate array model every computational step is viewed as being quantum. The notion of classical-quantum separation becomes more compelling when we consider say, Shor's algorithm in its full totality, including the significant amount of non-trivial classical post-processing of measurement results needed to reach the final answer. It seems inappropriate to view this post-processing as a quantum process (albeit one that maintains the computational basis)!

In measurement based computation the quantum parts of the algorithm are the quantum measurements done in parallel layers and the interspersed classical parts correspond to the adaptive choices of measurement bases, determined by classical computations on the previous layers' measurement outcomes. We may generalise this formalism in the following way: we allow (adaptively chosen) unitary gates as well as measurements within the quantum parts. We allow quantum layers to have only depth 1 (so a depth K quantum process is regarded as K layers with no interspersed classical computations) whereas classical layers can have any depth i.e. we are less concerned about controlling their structure.

In this formalism any quantum computation is viewed as a sequence of classical and quantum layers. The total quantum state is passed from one quantum layer to the next and the quantum actions carried out in the next layer are determined by classical computations on measurement outcomes from previous layers.

Any polynomial time quantum computation (say in the gate array model) can clearly be implemented with a polynomial number of quantum layers (and no

interspersed classical layers) but the above formalism suggests a novel structural conjecture:

Conjecture: Any polynomial time quantum algorithm can be implemented with only $O(\log n)$ quantum layers interspersed with polynomial time classical computations. □

This conjecture, asserting an exponential reduction in the essential "quantum content" of any quantum algorithm, has no analogue in classical complexity theory (where there is no notion of classical-quantum separation). Intuitively we are conjecturing that polynomial time classical computation needs relatively little "quantum assistance" to achieve the full power of polynomial time quantum computation. Although the conjecture remains unproven in general, we note that Cleve and Watrous [19] have shown that it holds true for Shor's algorithm.

9. Acknowledgements

These notes were developed in the course of presentations of the subject matter at a series of summer schools and workshops: the CNRS summer school on quantum logic and communication, Corsica, August 2004, the workshop on quantum information and computation, Newton Institute, Cambridge UK, August – December 2004, the NATO-ASI summer school on quantum computation and quantum information, Crete, May 2005 and the JST summer school on quantum information, Kochi, Japan, August 2005. Thanks to Sean Clark and Noah Linden for discussions of the material and to Sean Clark for providing the figures. This work was partially supported by the EU project RESQ-IST-2001-37559 and the UK EPSRC Interdisciplinary Research Collaboration on Quantum Information Processing.

References

[1] C. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres and W. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, *Phys. Rev Lett.* **70**, 1895-1899, 1993.

[2] D. Gottesman and I. Chuang, Quantum teleportation as a universal computational primitive, *Nature* **402**, 390-393, 1999. arXiv:quant-ph/9908010.

[3] M. A. Nielsen. Quantum computation by measurement and quantum memory. *Phys. Lett. A*, **308**, 96–100, 2003. arXiv:quant-ph/0108020.

[4] D. W. Leung. Quantum computation by measurements, *Int. J. Quant. Inf.* **2**, 33-43, (2004) arXiv:quant-ph/0310189, 2003. D. W. Leung. Two qubit projective measurements are universal for quantum computation. arXiv:quant-ph/0111122, 2001.

[5] R. Raussendorf and H. J. Briegel, A one-way quantum computer. *Phys. Rev. Lett.*, **86**, 5188–5191, 2001. arXiv:quant-ph/0010033.

[6] R. Raussendorf, D. E. Browne, and H. J. Briegel. Measurement-based quantum computation with cluster states *Phys. Rev. A*, **68**, 022312, 2003. arXiv:quant-ph/0301052.

[7] M. A. Nielsen and C. M. Dawson. Fault-tolerant quantum computation with cluster states. arXiv:quant-ph/0405134, 2004.

[8] R. F. Werner, All teleportation and dense coding schemes, arXiv:quant-ph/0003070, 2000.
[9] D. Gottesman, Stabilizer Codes and Quantum Error Correction, PhD thesis, California Institute of Technology, Pasadena, CA, 1997. arXiv:quant-ph/9705052. See also D. Gottesman, Course on quantum error correction, Perimeter Institute, Waterloo http://perimeterinstitute.ca/people/researchers/dgottesman/CO639-2004/index.html
[10] M. A. Nielsen and I. Chuang, Quantum Computation and Information, CUP 2000, page 464.
[11] S. Aaronson and D. Gottesman, Improved simulation of stabiliser circuits. arXiv:quant-ph/0406196, 2004.
[12] R. Raussendorf and H. J.Briegel, Computational model underlying the one-way quantum computer. *Quant. Inf. Comp.* **2**, 433, 2002. arXiv:quant-ph/0108067, 2001.
[13] Yaoyun Shi, Both Toffoli and Controlled-NOT need little help to do universal quantum computation. arXiv:quant-ph/0205115, 2002.
[14] M. A. Nielsen, Cluster-state quantum computation, to appear in *Rev. Math. Phys.* arXiv:quant-ph/0504097, 2005.
[15] P. Aliferis and D. W. Leung, Computation by measurements: a unifying picture. arXiv:quant-ph/0404082, 2004.
[16] A. M. Childs, D. W. Leung, and M. A. Nielsen, Unified derivation of measurement-based schemes for quantum computation. arXiv:quant-ph/0404132, 2004.
[17] P. Jorrand and S. Perdrix, Unifying quantum computation with projective measurements only and one-way quantum computation. arXiv:quant-ph/0404125, 2004.
[18] F. Verstraete and J. I. Cirac, Valence bond solids for quantum computation. *Phys. Rev. A*, **70**, 060302 (R), 2004. arXiv:quant-ph/0311130.
[19] R. Cleve and J. Watrous, fast parallel circuits for the quantum Fourier transform, Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science (FOCS 2000), 526-536, 2000. arXiv:quant-ph/0006004, 2000.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Quantum Error Correction and Fault-Tolerance

Daniel Gottesman
Perimeter Institute, Waterloo, Ontario, Canada

Abstract. I give an overview of the basic concepts behind quantum error correction and quantum fault tolerance. This includes the quantum error correction conditions, stabilizer codes, CSS codes, transversal gates, fault-tolerant error correction, and the threshold theorem.

Keywords. quantum computing, quantum error correction, fault-tolerant quantum computation

1. Quantum Error Correction

Building a quantum computer or a quantum communications device in the real world means having to deal with errors. Any qubit stored unprotected or one transmitted through a communications channel will inevitably come out at least slightly changed. The theory of quantum error-correcting codes has been developed to counteract noise introduced in this way. By adding extra qubits and carefully encoding the quantum state we wish to protect, a quantum system can be insulated to great extent against errors.

To build a quantum computer, we face an even more daunting task: If our quantum gates are imperfect, everything we do will add to the error. The theory of fault-tolerant quantum computation tells us how to perform operations on states encoded in a quantum error-correcting code without compromising the code's ability to protect against errors. Quantum error correction has other applications, as well. Quantum codes contain interesting entangled states, and techniques used to study quantum codes can tell us about entanglement. Quantum error correction has also been used to study quantum cryptography, for instance to prove the security of quantum key distribution.

In general, a quantum error-correcting code is a subspace of a Hilbert space designed so that any of a set of possible errors can be corrected by an appropriate quantum operation. Specifically:

Definition 1 *Let $\mathcal{H}_n$ be a 2^n-dimensional Hilbert space (n qubits), and let C be a K-dimensional subspace of $\mathcal{H}_n$. Then C is an $((n, K))$ (binary) quantum error-correcting code (QECC) correcting the set of errors $\mathcal{E} = \{E_a\}$ iff $\exists \mathcal{R}$ s.t. $\mathcal{R}$ is a quantum operation and $(\mathcal{R} \circ E_a)(|\psi\rangle) = |\psi\rangle$ for all $E_a \in \mathcal{E}$, $|\psi\rangle \in C$.*

$\mathcal{R}$ is called the *recovery* or *decoding* operation and serves to actually perform the correction of the state. The decoder is sometimes also taken to map $\mathcal{H}_n$ into

an unencoded Hilbert space $\mathcal{H}_{\log K}$ isomorphic to C. This should be distinguished from the *encoding* operation which maps $\mathcal{H}_{\log K}$ into $\mathcal{H}_n$, determining the imbedding of C. The computational complexity of the encoder is frequently a great deal less than that of the decoder. In particular, the task of determining what error has occurred can be computationally difficult (NP-hard, in fact), and designing codes with efficient decoding algorithms is an important task in quantum error correction, as in classical error correction.

This article will cover only binary quantum codes, built with qubits as registers, but all of the techniques discussed here can be generalized to higher-dimensional registers, or *qudits*. Qudit-based QECCs are useful not only for protecting systems built from physical qudits, but also can be used as building blocks to design better qubit-based codes.

To determine whether a given subspace is able to correct a given set of errors, we can apply the quantum error-correction conditions [2,7]:

Theorem 1 *A QECC C corrects the set of errors $\mathcal{E}$ iff*

$$\langle\psi_i|E_a^\dagger E_b|\psi_j\rangle = C_{ab}\delta_{ij}, \tag{1}$$

where $E_a, E_b \in \mathcal{E}$ and $\{|\psi_i\rangle\}$ form an orthonormal basis for C.

The salient point in these error-correction conditions is that the matrix element C_{ab} does not depend on the encoded basis states i and j, which roughly speaking indicates that neither the environment nor the decoding operation learns any information about the encoded state. We can imagine the various possible errors taking the subspace C into other subspaces of $\mathcal{H}_n$, and we want those subspaces to be isomorphic to C, and to be distinguishable from each other by an appropriate measurement. For instance, if $C_{ab} = \delta_{ab}$, then the various erroneous subspaces are orthogonal to each other.

Because of the linearity of quantum mechanics, we can always take the set of errors $\mathcal{E}$ to be a linear space: If a QECC corrects E_a and E_b, it will also correct $\alpha E_a + \beta E_b$ using the same recovery operation. In addition, if we write any superoperator $\mathcal{S}$ in terms of its operator-sum representation $\mathcal{S}(\rho) \mapsto \sum A_k \rho A_k^\dagger$, a QECC that corrects the set of errors $\{A_k\}$ automatically corrects $\mathcal{S}$ as well. Thus, it is sufficient in general to check that the error-correction conditions hold for a basis of errors.

Frequently, we are interested in codes that correct any error affecting t or fewer physical qubits. In that case, let us consider tensor products of the Pauli matrices

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2}$$

Define the Pauli group $\mathcal{P}_n$ as the group consisting of tensor products of I, X, Y, and Z on n qubits, with an overall phase of ± 1 or $\pm i$. The *weight* $\mathrm{wt}(P)$ of a Pauli operator $P \in \mathcal{P}_n$ is the number of qubits on which it acts as X, Y, or Z (i.e., not as the identity). Then the Pauli operators of weight t or less form a basis for the set of all errors acting on t or fewer qubits, so a QECC which corrects these Pauli operators corrects all errors acting on up to t qubits. If we have a

channel which causes errors independently with probability $O(\epsilon)$ on each qubit in the QECC, then the code will allow us to decode a correct state except with probability $O(\epsilon^{t+1})$, which is the probability of having more than t errors. We get a similar result in the case where the noise is a general quantum operation on each qubit which differs from the identity by something of size $O(\epsilon)$.

Definition 2 *The distance d of an $((n, K))$ QECC is the smallest weight of a nontrivial Pauli operator $E \in \mathcal{P}_n$ s.t. the equation*

$$\langle \psi_i | E | \psi_j \rangle = C(E)\delta_{ij} \tag{3}$$

fails.

We use the notation $((n, K, d))$ to refer to an $((n, K))$ QECC with distance d. Note that for $P, Q \in \mathcal{P}_n$, $\mathrm{wt}(PQ) \leq \mathrm{wt}(P) + \mathrm{wt}(Q)$. Then by comparing the definition of distance with the quantum error-correction conditions, we immediately see that a QECC corrects t general errors iff its distance $d > 2t$. If we are instead interested in *erasure* errors, when the location of the error is known but not its precise nature, a distance d code corrects $d - 1$ erasure errors. If we only wish to *detect* errors, a distance d code can detect errors on up to $d - 1$ qubits.

One of the central problems in the theory of quantum error correction is to find codes which maximize the ratios $(\log K)/n$ and d/n, so they can encode as many qubits as possible and correct as many errors as possible. Conversely, we are also interested in the problem of setting upper bounds on achievable values of $(\log K)/n$ and d/n. The quantum Singleton bound (or Knill-Laflamme bound [7]) states that any $((n, K, d))$ QECC must satisfy

$$n - \log K \geq 2d - 2. \tag{4}$$

Stronger bounds can be obtained using linear programming methods. We can set a lower bound on the existence of QECCs using the quantum Gilbert-Varshamov bound, which states that, for large n, an $((n, 2^k, d))$ QECC exists provided that

$$k/n \leq 1 - (d/n)\log 3 - h(d/n), \tag{5}$$

where $h(x) = -x \log x - (1 - x)\log(1 - x)$ is the binary Hamming entropy. Note that the Gilbert-Varshamov bound simply states that codes at least this good exist; it does not suggest that better codes cannot exist.

2. Stabilizer Codes

In order to better manipulate and discover quantum error-correcting codes, it is helpful to have a more detailed mathematical structure to work with. The most widely-used structure gives a class of codes known as *stabilizer codes* [3,5]. They are less general than arbitrary quantum codes, but have a number of useful properties that make them easier to work with than the general QECC.

Definition 3 *Let $S \subset \mathcal{P}_n$ be an Abelian subgroup of the Pauli group that does not contain -1 or $\pm i$, and let $C(S) = \{|\psi\rangle$ s.t. $P|\psi\rangle = |\psi\rangle \ \forall P \in S\}$. Then $C(S)$ is a stabilizer code and S is its* stabilizer.

Because of the simple structure of the Pauli group, any Abelian subgroup has order 2^{n-k} for some k and can easily be specified by giving a set of $n-k$ commuting generators.

The codewords of the QECC are by definition in the $+1$-eigenspace of all elements of the stabilizer, but an error E acting on a codeword will move the state into the -1-eigenspace of any stabilizer element M which anticommutes with E:

$$M\left(E|\psi\rangle\right) = -EM|\psi\rangle = -E|\psi\rangle. \tag{6}$$

Thus, measuring the eigenvalues of the generators of S tells us information about the error that has occurred. The set of such eigenvalues can be represented as an $(n-k)$-dimensional binary vector known as the *error syndrome* (0 for eigenvalue $+1$, 1 for eigenvalue -1). Note that the error syndrome does not tell us anything about the encoded state, only about the error that has occurred.

Theorem 2 *Let S be a stabilizer with $n-k$ generators, and let $S^{\perp} = \{E \in \mathcal{P}_n$ s.t. $[E,M] = 0\ \forall M \in S\}$. Then S encodes k qubits and has distance d, where d is the smallest weight of an operator in $S^{\perp} \setminus S$.*

We use the notation $[[n,k,d]]$ to a refer to such a stabilizer code. Note that the square brackets specify that the code is a stabilizer code, and that the middle term k refers to the number of encoded qubits, and not the dimension 2^k of the encoded subspace, as for the general QECC (whose dimension might not be a power of 2).

$S^{\perp}$ is the set of Pauli operators that commute with all elements of the stabilizer. They would therefore appear to be those errors which cannot be detected by the code. However, the theorem specifies the distance of the code by considering $S^{\perp} \setminus S$. (Since S is Abelian, $S \subseteq S^{\perp}$ always.) A Pauli operator $P \in S$ cannot be detected by the code, but there is in fact no need to detect it, since all codewords remain fixed under P, making it equivalent to the identity operation. A distance d stabilizer code which has nontrivial $P \in S$ with $\mathrm{wt}(P) < d$ is called *degenerate*, whereas one which does not is *non-degenerate*. A general non-stabilizer code is non-degenerate if the matrix C_{ab} in the error-correction conditions has maximum rank when the set of errors $\mathcal{E}$ is linearly independent. The phenomenon of degeneracy has no analogue for classical error correcting codes, and makes the study of quantum codes substantially more difficult than the study of classical error correction. For instance, a standard bound on classical error correction is the Hamming bound (or sphere-packing bound), but the analogous quantum Hamming bound

$$k/n \leq 1 - (t/n)\log 3 - h(t/n) \tag{7}$$

for $[[n,k,2t+1]]$ codes (when n is large) is only known to apply to non-degenerate quantum codes (though in fact we do not know of any degenerate QECCs that violate the quantum Hamming bound).

An example of a stabilizer code is the 5-qubit code, a $[[5,1,3]]$ code whose stabilizer can be generated by

$$X \otimes Z \otimes Z \otimes X \otimes I,$$
$$I \otimes X \otimes Z \otimes Z \otimes X,$$
$$X \otimes I \otimes X \otimes Z \otimes Z,$$
$$Z \otimes X \otimes I \otimes X \otimes Z.$$

The 5-qubit code is a non-degenerate code, and is the smallest possible QECC which corrects 1 error (as one can see from the quantum Singleton bound).

It is frequently useful to consider other representations of stabilizer codes. For instance, $P \in \mathcal{P}_n$ can be represented by a pair of n-bit binary vectors $(p_X|p_Z)$ where p_X is 1 for any location where P has an X or Y tensor factor and is 0 elsewhere, and p_Z is 1 for any location where P has a Y or Z tensor factor. Two Pauli operators $P = (p_X|p_Z)$ and $Q = (q_X|q_Z)$ commute iff $p_X \cdot q_Z + p_Z \cdot q_X = 0$. Then the stabilizer for a code becomes a pair of $(n-k) \times n$ binary matrices, and most interesting properties can be determined by an appropriate linear algebra exercise. Another useful representation is to map the single-qubit Pauli operators I, X, Y, Z to the finite field GF(4), which sets up a connection between stabilizer codes and a subset of classical codes on 4-dimensional registers.

3. CSS Codes

CSS codes are a very useful class of stabilizer codes invented by Calderbank and Shor, and by Steane [4,10]. The construction takes two binary classical linear codes and produces a quantum code, and can therefore take advantage of much existing knowledge from classical coding theory. In addition, CSS codes have some very useful properties which make them excellent choices for fault-tolerant quantum computation.

A classical $[n, k, d]$ linear code (n physical bits, k logical bits, classical distance d) can be defined in terms of an $(n-k) \times n$ binary *parity check* matrix H — every classical codeword v must satisfy $Hv = 0$. Each row of the parity check matrix can be converted into a Pauli operator by replacing each 0 with an I operator and each 1 with a Z operator. Then the stabilizer code generated by these operators is precisely a quantum version of the classical error-correcting code given by H. If the classical distance $d = 2t + 1$, the quantum code can correct t bit flip (X) errors, just as could the classical code.

If we want to make a QECC that can also correct phase (Z) errors, we should choose *two* classical codes C_1 and C_2, with parity check matrices H_1 and H_2. Let C_1 be an $[n, k_1, d_1]$ code and let C_2 be an $[n, k_2, d_2]$ code. We convert H_1 into stabilizer generators as above, replacing each 0 with I and each 1 with Z. For H_2, we perform the same procedure, but each 1 is instead replaced by X. The code will be able to correct bit flip (X) errors as if it had a distance d_1 and to correct phase (Z) errors as if it had a distance d_2. Since these two operations are completely separate, it can also correct Y errors as both a bit flip and a phase error. Thus, the distance of the quantum code is at least $\min(d_1, d_2)$, but might be higher because of the possibility of degeneracy.

However, in order to have a stabilizer code at all, the generators produced by the above procedure must commute. Define the dual $C^{\perp}$ of a classical code C as the set of vectors w s.t. $w \cdot v = 0$ for all $v \in C$. Then the Z generators from H_1 will all commute with the X generators from H_2 iff $C_2^{\perp} \subseteq C_1$ (or equivalently, $C_1^{\perp} \subseteq C_2$). When this is true, C_1 and C_2 define an $[[n, k_1 + k_2 - n, d]]$ stabilizer code, where $d \geq \min(d_1, d_2)$.

One useful property of CSS codes is that the quantum codewords are superpositions of classical codewords. Indeed, we can choose each basis codeword of the CSS code to be a superposition (with equal amplitude) over the words in some coset of $C_2^{\perp}$ inside of C_1. For example, the smallest distance-3 CSS code is the 7-qubit code, a $[[7, 1, 3]]$ QECC created from the classical Hamming code (consisting of all sums of classical strings 1111000, 1100110, 1010101, and 1111111). The encoded $|\overline{0}\rangle$ for this code consists of the superposition of all even-weight classical codewords and the encoded $|\overline{1}\rangle$ is the superposition of all odd-weight classical codewords. The 7-qubit code is much studied because its properties make it particularly well-suited to fault-tolerant quantum computation.

4. Fault-Tolerance

By themselves, the quantum error-correcting codes described above are useful for transmitting quantum states over a noisy quantum channel or storing quantum states for longer than the decoherence time. However, if we wish to go further and perform computations on an encoded state while it is actively experiencing noise, we must supplement the QECC with a protocol for fault-tolerant operations.

The basic design principle of a fault-tolerant protocol is that an error in a single location — either a faulty gate or noise on a quiescent qubit — should not be able to alter more than a single qubit in each block of the quantum error-correcting code. If this condition is satisfied, t separate single-qubit or single-gate failures are required for a distance $2t + 1$ code to fail.

Particular caution is necessary, as computational gates can cause errors to propagate from their original location onto qubits that were previously correct. In general, a gate coupling pairs of qubits allows errors to spread in both directions across the coupling.

It is easy to see an example of this with the CNOT gate: Imagine we were to perform the CNOT on the two qubits $|0\rangle\,|0\rangle$. The correct outcome is again $|0\rangle\,|0\rangle$. But now suppose that there is an error on the first qubit, so the actual state is $|1\rangle\,|0\rangle$; then the CNOT produces the state $|1\rangle\,|1\rangle$. Both qubits are now wrong.

In the quantum case, the presence of additional types of errors mean that the CNOT can propagate errors from the target qubit to the control qubit as well as from the control to the target, as occurs with the classical CNOT. For instance, imagine we wish to perform CNOT on $|+\rangle\,|+\rangle$, but instead perform it on $|+\rangle\,|-\rangle$ (where $|\pm\rangle = |0\rangle \pm |1\rangle$). The outcome is then $|-\rangle\,|-\rangle$ instead of the correct state $|+\rangle\,|+\rangle$ — the error has propagated backwards along the CNOT.

The solution is to use transversal gates whenever possible [9]. A transversal operation is one in which the ith qubit in each block of a QECC interacts only with the ith qubit of other blocks of the code or of special ancilla states. An

operation consisting only of single-qubit gates is automatically transversal. A transversal operation has the virtue that an error occurring on the 3rd qubit in a block, say, can only ever propagate to the 3rd qubit of other blocks of the code, no matter what other sequence of gates we perform before a complete error-correction procedure.

In the case of certain codes, such as the 7-qubit code, a number of different gates can be performed transversally. Unfortunately, it does not appear to be possible to perform universal quantum computations using just transversal gates. We therefore have to resort to more complicated techniques. First we create special encoded ancilla states in a non-fault-tolerant way, but perform some sort of check on them (in addition to error correction) to make sure they are not too far off from the goal. Then we interact the ancilla with the encoded data qubits using gates from our stock of transversal gates and perform a fault-tolerant measurement. Then we complete the operation with a further transversal gate which depends on the outcome of the measurement.

5. Fault-Tolerant Gates

We will focus on stabilizer codes. Universal fault-tolerance is known to be possible for any stabilizer code, but in most cases the more complicated type of construction is needed for all but a few gates. The Pauli group $\mathcal{P}_k$, however, can be performed transversally on any stabilizer code. Indeed, the set $S^\perp \setminus S$ of undetectable errors is a boon in this case, as it allows us to perform these gates. In particular, each coset $S^\perp/S$ corresponds to a different logical Pauli operator (with S itself corresponding to the identity). On a stabilizer code, therefore, logical Pauli operations can be performed via a transversal Pauli operation on the physical qubits.

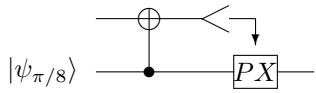
Stabilizer codes have a special relationship to a finite subgroup $\mathcal{C}_n$ of the unitary group $U(2^n)$ frequently called the *Clifford group*. The Clifford group on n qubits is defined as the set of unitary operations which conjugate the Pauli group $\mathcal{P}_n$ into itself; $\mathcal{C}_n$ can be generated by the Hadamard transform, the CNOT, and the single-qubit $\pi/4$ phase rotation $\mathrm{diag}(1, i)$. The set of stabilizer codes is exactly the set of codes which can be created by a Clifford group encoder circuit using $|0\rangle$ ancilla states.

Some stabilizer codes have interesting symmetries under the action of certain Clifford group elements, and these symmetries result in transversal gate operations. A particularly useful fact is that a transversal CNOT gate (i.e., CNOT acting between the ith qubit of one block of the QECC and the ith qubit of a second block for all i) acts as a logical CNOT gate on the encoded qubits for any CSS code. Furthermore, for the 7-qubit code, transversal Hadamard performs a logical Hadamard, and the transversal $\pi/4$ rotation performs a logical $-\pi/4$ rotation. Thus, for the 7-qubit code, the full logical Clifford group is accessible via transversal operations.

Unfortunately, the Clifford group by itself does not have much computational power: it can be efficiently simulated on a classical computer. We need to add some additional gate outside the Clifford group to allow universal quantum com-

putation; a single gate will suffice. Two common choices for the additional gate are the single-qubit $\pi/8$ phase rotation $\mathrm{diag}(1, \exp(i\pi/4))$ or the three-qubit Toffoli gate (controlled-controlled-NOT). Note that either of these gates plus the Clifford group comprise a finite generating set of gates. However, by taking appropriate products, we get an infinite set of gates, one that is dense in the unitary group $U(2^n)$, allowing universal quantum computation.

The following circuit performs a $\pi/8$ rotation, given an ancilla state $|\psi_{\pi/8}\rangle = |0\rangle + \exp(i\pi/4)|1\rangle$:



Here P is the $\pi/4$ phase rotation $\mathrm{diag}(1, i)$, and X is the bit flip. The product is in the Clifford group, and is only performed if the measurement outcome is 1. Therefore, given the ability to perform fault-tolerant Clifford group operations, fault-tolerant measurements, and to prepare the encoded $|\psi_{\pi/8}\rangle$ state, we have universal fault-tolerant quantum computation. A slight generalization of the fault-tolerant measurement procedure below can be used to fault-tolerantly verify the $|\psi_{\pi/8}\rangle$ state, which is a $+1$ eigenstate of PX. Using this or another verification procedure, we can check a non-fault-tolerant construction or simply measure the eigenvalue of PX on a random state (the -1 eigenstate differs from the $+1$ eigenstate simply by a phase flip Z).

6. Fault-Tolerant Measurement and Error Correction

Since all our gates are unreliable, including those used to correct errors, we will need some sort of fault-tolerant quantum error correction procedure. A number of different techniques have been developed. All of them share some basic features: they involve creation and verification of specialized ancilla states, and use transversal gates which interact the data block with the ancilla state.

The simplest method, due to Shor, is very general but also requires the most overhead and is frequently the most susceptible to noise. Note that the following procedure can be used to measure (non-fault-tolerantly) the eigenvalue of any (possibly multi-qubit) Pauli operator M: Produce an ancilla qubit in the state $|+\rangle = |0\rangle + |1\rangle$. Perform a controlled-$M$ operation from the ancilla to the state being measured. In the case where M is a multi-qubit Pauli operator, this can be broken down into a sequence of controlled-X, controlled-Y, and controlled-Z operations. Then measure the ancilla in the basis of $|+\rangle$ and $|-\rangle = |0\rangle - |1\rangle$. If the state is a $+1$ eigenvector of M, the ancilla will be $|+\rangle$, and if the state is a -1 eigenvector, the ancilla will be $|-\rangle$.

The advantage of this procedure is that it measures just M and nothing more. The disadvantage is that it is not transversal, and thus not fault-tolerant. Instead of the unencoded $|+\rangle$ state, we must use a more complex ancilla state

$|00\ldots0\rangle+|11\ldots1\rangle$ known as a "cat" state (after Schrödinger's cat). The cat state contains as many qubits as the operator M to be measured, and we perform the controlled-X, -Y, or -Z operations transversally from the appropriate qubits of the cat state to the appropriate qubits in the data block. Since, assuming the cat state is correct, all of its qubits are either $|0\rangle$ or $|1\rangle$, the procedure either leaves the data state alone or performs M on it uniformly. A $+1$ eigenstate in the data therefore leaves us with $|00\ldots0\rangle + |11\ldots1\rangle$ in the ancilla and a -1 eigenstate leaves us with $|00\ldots0\rangle - |11\ldots1\rangle$. In either case, the final state still tells us nothing about the data beyond the eigenvalue of M. If we perform a Hadamard transform and then measure each qubit in the ancilla, we get either a random even weight string (for eigenvalue $+1$) or an odd weight string (for eigenvalue -1).

The procedure is transversal, so an error on a single qubit in the initial cat state or in a single gate during the interaction will only produce one error in the data. However, the initial construction of the cat state is not fault-tolerant, so a single gate error then could eventually produce two errors in the data block. Therefore, we must be careful and use some sort of technique to verify the cat state, for instance by checking if random pairs of qubits are the same. Also, note that a single phase error in the cat state will cause the final measurement outcome to be wrong (even and odd switch places), so we should repeat the measurement procedure multiple times for greater reliability.

We can then make a full fault-tolerant error correction procedure by performing the above measurement technique for each generator of the stabilizer. Each measurement gives us one bit of the error syndrome, which we then decipher classically to determine the actual error.

More sophisticated techniques for fault-tolerant error correction involve less interaction with the data but at the cost of more complicated ancilla states. A procedure due to Steane uses (for CSS codes) one ancilla in a logical $|\overline{0}\rangle$ state of the same code and one ancilla in a logical $|\overline{0}\rangle+|\overline{1}\rangle$ state. A procedure due to Knill (for any stabilizer code) teleports the data qubit through an ancilla consisting of two blocks of the QECC containing an encoded Bell state $|\overline{00}\rangle+|\overline{11}\rangle$. Because the ancillas in Steane and Knill error correction are more complicated than the cat state, it is especially important to verify the ancillas before using them.

7. Threshold for Fault-Tolerance

In an unencoded protocol, even one error can destroy the computation, but a fully fault-tolerant protocol will give the right answer unless multiple errors occur before they can be corrected. On the other hand, the fault-tolerant protocol is larger, requiring more qubits and more time to do each operation, and therefore providing more opportunities for errors. If errors occur on the physical qubits independently at random with probability p per gate or timestep, the fault-tolerant protocol has probability of logical error for a single logical gate or timestep at most Cp^2, where C is a constant that depends on the design of the fault-tolerant circuitry (assume the quantum error-correcting code has distance 3, as for the 7-qubit code). When $p < p_t = 1/C$, the fault-tolerance helps, decreasing the log-

ical error rate. p_t is the *threshold* for fault-tolerant quantum computation. If the error rate is higher than the threshold, the extra overhead means that errors will occur faster than they can be reliably corrected, and we are better off with an unencoded system.

To further lower the logical error rate, we turn to a family of codes known as *concatenated codes* [1,6,8]. Given a codeword of a particular $[[n, 1]]$ QECC, we can take each physical qubit and again encode it using the same code, producing an $[[n^2, 1]]$ QECC. We could repeat this procedure to get an n^3-qubit code, and so forth. The fault-tolerant procedures concatenate as well, and after L levels of concatenation, the effective logical error rate is $p_t(p/p_t)^{2^L}$ (for a base code correcting 1 error). Therefore, if p is below the threshold p_t, we can achieve an arbitrarily good error rate ϵ per logical gate or timestep using only $\mathrm{poly}(\log \epsilon)$ resources, which is excellent theoretical scaling.

Unfortunately, the practical requirements for this result are not nearly so good. The best rigorous proofs of the threshold to date show that the threshold is at least 2×10^{-5} (meaning one error per 50,000 operations), which is challenging, but may be achievable. Optimized simulations of fault-tolerant protocols suggest the true threshold may be as high as 5%, but to tolerate this much error, existing protocols require enormous overhead, perhaps increasing the number of gates and qubits by a factor of a million or more for typical computations. For lower physical error rates, overhead requirements are more modest, particularly if we only attempt to optimize for calculations of a given size, but are still larger than one would like.

Furthermore, these calculations make a number of assumptions about the physical properties of the computer. The errors are assumed to be independent and uncorrelated between qubits except when a gate connects them. It is assumed that measurements and classical computations can be performed quickly and reliably, and that quantum gates can be performed between arbitrary pairs of qubits in the computer, irrespective of their physical proximity. Of these, only the assumption of independent errors is at all necessary, and that can be considerably relaxed to allow short-range correlations and certain kinds of non-Markovian environments. However, the effects of relaxing these assumptions on the threshold value and overhead requirements have not been well-studied.

References

[1] D. Aharonov and M. Ben-Or, "Fault-tolerant quantum computation with constant error rate," quant-ph/9906129.

[2] C. Bennett, D. DiVincenzo, J. Smolin, and W. Wootters, "Mixed state entanglement and quantum error correction," Phys. Rev. A **54** (1996), 3824–3851; quant-ph/9604024.

[3] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "Quantum error correction via codes over GF(4)," IEEE Trans. Inform. Theory **44** (1998), 1369–1387; quant-ph/9605005.

[4] A. R. Calderbank and P. W. Shor, "Good quantum error-correcting codes exist," Phys. Rev. A **54** (1996), 1098–1105; quant-ph/9512032.

[5] D. Gottesman, "Class of quantum error-correcting codes saturating the quantum Hamming bound," Phys. Rev. A **54** (1996), 1862–1868; quant-ph/9604038.

[6] A. Y. Kitaev, "Quantum error correction with imperfect gates," Quantum Communication, Computing, and Measurement (Proc. 3rd Int. Conf. of Quantum Communication and Measurement) (Plenum Press, New York, 1997), p. 181–188.
[7] E. Knill and R. Laflamme, "A theory of quantum error-correcting codes," Phys. Rev. A **55** (1997), 900–911; quant-ph/9604034.
[8] E. Knill, R. Laflamme, and W. H. Zurek, "Resilient quantum computation," Science **279** (1998), 342–345.
[9] P. W. Shor, "Fault-tolerant quantum computation," Proc. 35th Ann. Symp. on Fundamentals of Computer Science (IEEE Press, Los Alamitos, 1996), pp. 56–65; quant-ph/9605011.
[10] A. M. Steane, "Multiple particle interference and quantum error correction," Proc. Roy. Soc. London A **452** (1996), 2551–2577; quant-ph/9601029.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Simulating Fourier Transforms for Open Quantum Systems in Higher Encoding Bases

I. N. Doxaras [a,1]
[a] *Technical University of Athens*
Division of Theoretical Physics

Abstract. We describe the formalism of gate level simulation in higher dimensional encoding systems. Furthermore, we introduce efficient error–correcting codes for such systems and use coding theory to test the error prevention in prime bases. By making extensive use of quantum Fourier transforms, we simulate open quantum systems and benchmark the efficiency of the error–correcting algorithm using the QCHBS package [1]. Numerical simulations indicate greater stability of prime base encoding systems.

Keywords. qutrits, error–correction, Quantum Fourier transform, simulation, qudits

1. The Formalism of Encoding in Higher Dimensions

The extensive theoretical exploration of higher encoding systems is justified by the clear advantages that such systems exhibit. A significant body of evidence leads us to believe that they obtain a remarkable resistance to quantum noise [2,3,4,5], produce more efficient cryptographic schemes [6] and generate quantum mechanical states with enhanced entanglement [7,8], which is the goal of every efficient implementation of a quantum computer. Furthermore, using such encoding bases, we experience a contraction of the quantum register by a factor $\log_d N$. Such a reduction in the number of entangled qudits has the significant impact of increasing the decoherence times for specific implementations. The generalisation from the qubit formalism is straight forward, considering the mathematical constraints that the Hilbert space imposes on our construction. Every state is defined using the basis vectors of the space $\mathbb{H}_d$ as follows $|k\rangle = (0, 0, \ldots, \underbrace{1}_{k}, 0, \ldots, 0)^T$.

We use the notion of the generalised Pauli group [11] to denote all the possible representations of the $SU(d)$ group, the group of all complex unitary matrices with unit trace. These constitute the set of all possible transforms in a d–dimensional Hilbert space. Generalising also the concept of the Bloch sphere representation, we interpret this group

[1]Correspondence to: I. N. Doxaras, Tel.: +30 210 7721712; National Technical University of Athens, Department of Physics and Applied Mathematics, E-mail: doxaras@inp.demokritos.gr.

as the complete set of rotations on a d-dimensional Bloch hypersphere. Some elementary gates with great significance in our further analysis are

$$Phase_d\,|j\rangle \mapsto \omega^j|j\rangle; \qquad S_d\,|j\rangle \mapsto |d\,j\rangle; \qquad SUM\,|j\rangle|k\rangle \mapsto |j\rangle|j +_d k\rangle \tag{1}$$

where $\omega = e^{2\pi\, i/d}$ and $+_d$ indicates addition modulo d.

Using simple algebra, it is also possible to restate the general Fourier transform [9] in $\mathbb{H}_d$ as a series of tensor products of unitary transforms (gates) applied iteratively on the initial state of our register. To introduce entanglement in our model we will use a Fourier transform matrix in the dimension d we consider. This is equivalent to the Hadamard gate when $d = 2$. The further use of Hadamards is limited by the Hadamard Conjecture[2]

$$\begin{aligned} |j\rangle &= \frac{1}{d^{n/2}} \sum_{k=0}^{d^n-1} e^{2\pi\, ijk/d^n}|k\rangle \\ &= \frac{1}{d^{n/2}} \bigotimes_{l=1}^{n} [|0\rangle + e^{2\pi ijd^{-l}}|1\rangle + \ldots (e^{2\pi\, ijd^{-l}})^{d-1}|d\rangle] \end{aligned} \tag{2}$$

2. Error Correction in Higher Dimensional Bases

Generalizations of error–correcting schemes and the stabilizer formalism are well known [10]. Using an algorithmic procedure, we are able to generate the stabiliser group, in order to achieve automatic encoding into the stabilizer formalism. This model is based on the cyclic property of the group $SU(d)$.

$$X_d = \begin{pmatrix} 0 & 0 & 0 & \ldots 1 \\ 1 & 0 & 0 & \ldots 0 \\ 0 & 1 & 0 & \ldots 0 \\ 0 & 0 & 1 & \ldots 0 \\ \ldots & \ldots & \ldots & \ldots 0 \end{pmatrix} \qquad Z_d = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & \omega^2 & 0 & \ldots & 0 \\ 0 & 0 & \omega^2 & \ldots & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & \ldots & \omega^{d-1} \end{pmatrix} \tag{3}$$

1. Initially we determine the irreducible representations of the group $SU(d)$ making use of the predefined matrices X_d and Z_d, which are written explicitly above. The key concepts to this construction are the closure of the group and that all the elements form a ring isomorphic to $\mathbb{Z}_d$ for d prime.
2. We iterate the product of these matrices using the fact that the group is cyclic and so, after at most d^2-3 iterations, all the possible d^2-1 irreducible representations of the generalised Pauli group will be produced. The algorithmic procedure is

$$\begin{aligned} &\text{for } i = 1 \text{ to } d; \\ &\quad \text{for } j = 1 \text{ to } d; \\ &\qquad\quad j < i \end{aligned}$$

[2]The Hadamard Conjecture states that such matrices exist only in 1, 2 and $4k$ dimensions, for integer k.

$$X_d^i \cdot Z_d^j$$

Introducing noise in the quantum channel is easy, after we have constructed the error group. Using a stochastic procedure, every error gate acts with a predefined probability p, corresponding to the unit transform I (preserved qubit) or $1-p$ when an interaction (error) with the environment occurs.

3. Simulation of the Error–Correcting Procedure

The theoretical background we used in our approach was described in the previous section. The simulator operates at the gate level and after acting with a specified gate of the quantum circuit, randomly selects an error gate which acts instantly after.

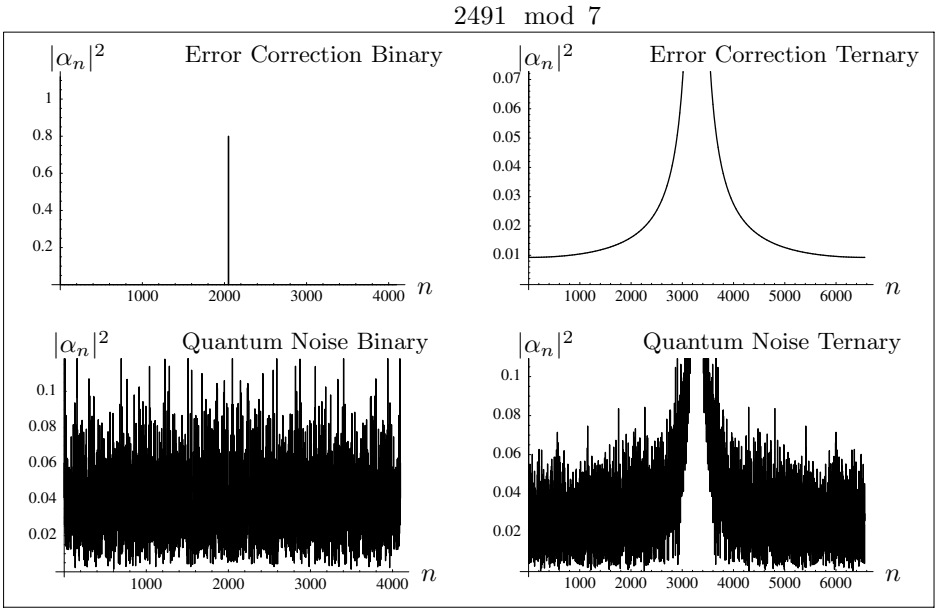


Figure 1. Output of Phase Detection Algorithm. This procedure is part of the quantum mechanical solution to the hidden subgroup problem when the group under consideration is Abelian. **1a.** The output of the closed quantum system encoded in binary base. Were we have ploted the probability of the outcome, versus the eigenvectors of $\mathbb{H}_{2^{\lfloor \log_2 2491+1 \rfloor}}$ space for qubits and $\mathbb{H}_{3^{\lfloor \log_3 2491+1 \rfloor}}$ space for qutrits. **1b.** The same output for the quantum phase estimation in a qutrit encoding system. **1c.** The quantum system upon interaction with the environment (thermal bath) is obviously distorted. To simulate environment interaction we operate with the error gates stochastically. We comment that the output does not have any retrievable information, it features white noise behaviour. **1d.** The output histogram indicates the existence of retrievable information achieved with qutrit encoding. The same behaviour is observed in higher bases like ququints and, in general, *qupits*, were the "p" stands for prime base.

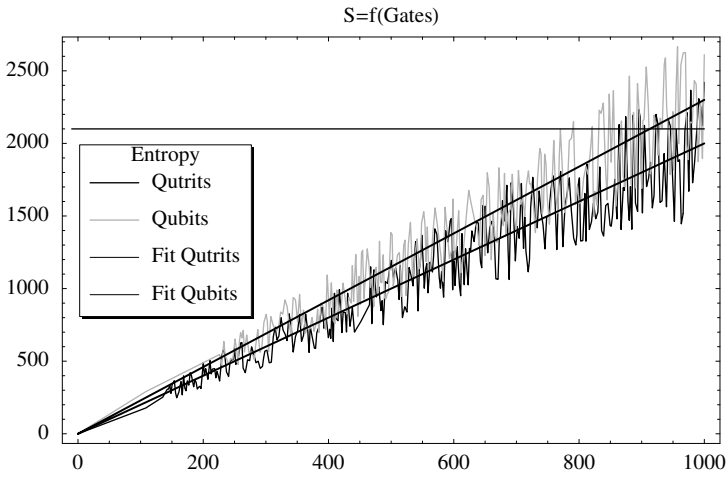


Figure 2. As an indication of information preservation for the quantum register we store a binary–ternary number in it and allow interaction with the environment; we plot the Von Neuman entropy ($S(\rho) \equiv -tr(\rho \log \rho)$) for the evolution of the Quantum register in $\mathbb{H}_d$ as related to the gates acted on. The Holevo bound for a Markovian stochastic procedure can be used as the upper bound of retrievable information in quantum channel, $H(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x)$. The bound is indicated on the plot. Our numerical results clearly indicate that in the simplest case of qutrits versus qubits, we enjoy higher error resistance and also higher information capacity when it comes to the quantum channel.

Acknowledgements

The author would like to thank Pr. E. G. Floratos of the University of Athens and Dr. M. Axenides of the NCSR Demokritos for inspiring discussions and warm hospitality at the institute of Nuclear Physics.

References

[1] *Quantum Computer Higher Base Simulator*
http://www.cc.uoa.gr/ MathPhys/qchbs/intro.html.
During the compilation of this paper we discovered the existence of a very similar simulator named *CHP*. S. Aaronson, D. Gottesman *Improved Simulation of Stabilizer Circuits*, arXiv:quant-ph/0406196.

[2] C. Brukner, M. Zukowski and A. Zeilinger, *Quantum communication complexity protocol with two entangled qutrits* Phys. Rev. Lett. **89** 197901, (2002).

[3] N. K. Langford et al. *Measuring Entangled Qutrits and Their Use for Quantum Bit Commitment* Phys. Rev. Lett. **93**, 053601 (2004)

[4] L-B Fu , J-L Chen and X-G Zhao, *Degree of entanglement for two qutrits*, Phys. Rev. A **68** 022323.
[5] H. Barnum, E. Knill, G. Ortiz, R. Somma and L. Viola *A subsystem-independent generalization of entanglement*, Phys. Rev. Lett. **92** 107902.
[6] D. Bruss, C.Macchiavello, *Optimal Eavesdropping in Cryptography with Three-Dimensional Quantum States*, Phys. Rev. Lett. 88, 127901 (2002).
[7] A. Grudka and A. Wòjcik, *How to encode the states of two non-entangled qubits in one qutrit*, Phys. Lett. A **314** 350.
[8] Y. Bogdanov et al., *Preparation of arbitrary qutrit state based on biphotons*, Opto-Ireland 2005, 2005SPIE.5833..202B.
[9] E.G. Floratos, Athanasiou and S. Nicolis, *Fast Quantum Maps*, arXiv:math-ph/9805012, (1998).
[10] D. Gottesman, PhD Thesis, *Stabiliser Codes and Quantum Error Correction*, arXiv:quant-ph/9705052
[11] A. Ashikhmin, E. Knill *Nonbinary Quantum Stabiliser Codes*, arXiv:quant-ph/0005008

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Entanglement, Area Law and Group Theory

Radu Ionicioiu [a] [1], Alioscia Hamma [a] and Paolo Zanardi [a]
[a] *Quantum Information Group, Institute for Scientific Interchange (ISI), Viale Settimio Severo 65, I-10133 Torino, Italy*

Abstract.
We investigate bipartite entanglement in a class of multiparty quantum states, the G-states, constructed from a group G, and we derive an expression for the von Neumann entropy. We show that for a special subset of such states, the G-homogeneous states, the entropy satisfies the area law. If G is a group of spin-flips, the G-homogeneous states are locally equivalent to 2-colorable graph states (e.g., n-GHZ, cluster states, etc). The advantage of our representation is a more compact description of these states in terms of a group of spin-flips G. As an example, we compute the n-tangle τ_n for 2-colorable graph states.

Keywords. Bipartite entanglement; stabilizer/graph states; area law.

1. Introduction

Entanglement emerged recently as an important tool in the study of several fields, including quantum information processing, lattice spin systems and quantum phase transitions. Despite the considerable effort put into classifying and quantifying entanglement for many-body systems, there is no known measure which completely characterizes the entanglement properties of an arbitrary system.

In this article we investigate bipartite entanglement using a group theoretical framework, giving an overview of the results and methods discussed in [1,2]. The main role will be played by a set of quantum states, the G-states, defined in terms of a group G acting on the Hilbert space of the system. We derive a general expression for the von Neumann entropy corresponding to a bipartition (A, B) of the full system. For a special class of states, the G-homogeneous states, we show that the entropy satisfies the area law. If G is a group of spin-flips, the G-homogeneous states are locally equivalent to a well-known class of entangled states, the 2-colorable graph states, which include n-GHZ and cluster states. Finally, using this new representation we obtain a simple formula for the n-tangle τ_n of an arbitrary 2-colorable graph state.

2. Entanglement, G-States and G-Homogeneous States

Let $\mathcal{H}$ be the Hilbert space of a quantum many-body system. In order to study bipartite entanglement in this system, we assume that we can partition it into two subsystems,

A and B, such that $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. In the case of a spin system, this corresponds to take A as a particular set of spins and B the complementary set. Our main tool will be the *entanglement entropy* (von Neumann entropy of subsystem A) defined as $S = -\mathrm{Tr}\,(\rho_A \log_2 \rho_A)$, where ρ_A is the reduced density matrix of the A subsystem.

The focus will be on a class of quantum states defined in terms of a group. Let G be a group having a bilocal action on $\mathcal{H}$, i.e., $\forall g \in G,\ \ g = g_A \otimes g_B$, where $g_{A,B}$ acts only on $\mathcal{H}_{A,B}$, respectively. Let $|0\rangle = |0_A\rangle \otimes |0_B\rangle \in \mathcal{H}$ be a reference product state; for a n-spins system we can take the reference state to be the fully polarized state $|0\rangle = |0\rangle_1 \otimes \ldots \otimes |0\rangle_n$. With these premises, we introduce a class of (normalized) states, the G-states, defined as

$$|\Psi_G\rangle := \sum_{g\in G} \alpha(g) g|0\rangle \tag{1}$$

A particularly important subset of the above class are the G-homogeneous states, for which all the coefficients are equal:

$$|G\rangle := \mathcal{N}^{-1/2} \sum_{g\in G} g|0\rangle \tag{2}$$

where $\mathcal{N} \neq 0$ is a normalizing factor. We point out an important property of G-homogeneous states, namely that they are *stabilized* by the group G, i.e., $g|G\rangle = |G\rangle,\ \forall g \in G$.

Consider now the two subgroups of G that act exclusively on the subsystems A and respectively B (the "local" groups):

$$G_A := \{g \in G \mid\ g = g_A \otimes \mathbb{1}_B\} \tag{3}$$

$$G_B := \{g \in G \mid\ g = \mathbb{1}_A \otimes g_B\} \tag{4}$$

As these subgroups are *normal*, we can define the quotient group (the "non-local" group)

$$G_{AB} := \frac{G}{G_A \times G_B} \tag{5}$$

This shows that we can partition G in equivalence classes $[h] \in G_{AB}$ as

$$G = \bigcup_{[h]\in G_{AB}} \{(g_A \otimes g_B)h \mid g_A \otimes \mathbb{1} \in G_A,\ \mathbb{1} \otimes g_B \in G_B\} \tag{6}$$

and inside each equivalence class the elements are parametrized by elements of the "local" group $G_A \times G_B$.

Theorem. Suppose we have a quantum system described by a G-state (1). Assume that the following separability condition holds for the coefficients $\alpha(g)$

$$\forall g \in G, \quad \alpha(g) \equiv \alpha(g_A \otimes g_B h) = \alpha_A(g_A)\alpha_B(g_B)\beta(h) \tag{7}$$

where $g_A \otimes \mathbb{1} \in G_A,\ \mathbb{1} \otimes g_B \in G_B$ and $[h] \in G_{AB}$; without loss of generality, we can take $\beta(h)$ to be normalized, $\sum_h |\beta(h)|^2 = 1$. Then the von Neumann entropy of the G-state corresponding to the bipartition (A, B) is bounded by

$$S(|\Psi_G\rangle) \leq -\sum_{[h]\in G_{AB}} |\beta(h)|^2 \log_2 |\beta(h)|^2 \tag{8}$$

For a proof we refer the reader to [1].
Corollary 1. The von Neumann entropy for a G-homogeneous state satisfies

$$S(|G\rangle) \leq \log_2 |G_{AB}| \tag{9}$$

Observation 1. The bounds in eqs.(8),(9) are saturated if the projections on the A and B subsystems of all (nontrivial) group elements (i.e., $g_{A,B}$) have zero expectation value on the reference state [1]:

$$g_x \neq \mathbb{1}_x \Rightarrow \langle 0_x|g_x|0_x\rangle = 0, \qquad x = A, B \tag{10}$$

We now discuss the significance of this result. The "nonlocal" group G_{AB} is generated by operators acting on both subsystems which cannot be factorized in unitaries of G acting only on A or B subsystems. Suppose we have a multi-partite quantum system whose degrees of freedom have nearest neighbor interaction and let A be the set of all the degrees of freedom inside a closed surface. Then the "local" groups G_A and G_B correspond to the bulk degrees of freedom in the two subsystems, whereas G_{AB} corresponds to the degrees of freedom localized on the boundary between A and B. Hence we can interpret eq. (9) as another manifestation of the so-called area law (or more generally, the Holographic Principle): *the entropy of a system is proportional to the boundary degrees of freedom and not to the bulk ones*. The generic behaviour expressed by the area law has been recovered in several (apparently unconnected) systems, like black holes, scalar fields on a lattice [3] and harmonic oscillators [4], to name only a few. For a group of spin flips (see next section), we have $|G_{AB}| = 2^{n_{AB}}$, where n_{AB} is the number of independent generators of G_{AB}. Then from eq.(9) the entropy is $S = n_{AB} \leq \Sigma$ and is bounded by the area Σ of the surface; here we can take as a measure of area the number of "punctures" a generator makes in the boundary, and hence $n_{AB} \leq \Sigma$.

From eq.(8) we can derive immediately a sufficient condition for $|\Psi_G\rangle$ to be separable with respect to the (A, B) partition.
Corollary 2. If the group is a direct product $G = G_A \times G_B$, then $S(|\Psi_G\rangle) = 0$.

3. G-Homogeneous States for a Group of Spin Flips

In the formalism discussed so far, G was an arbitrary, possibly non-Abelian, group having a bilocal action on $\mathcal{H}$. We now turn to a case of special interest, namely we take G to be a group of spin flips. In the following we will use the notation X_i, Y_i, Z_i for the (generalized) Pauli operators acting on the i-qubit (qudit). First we notice that an arbitrary state of a n-qubit system $|\psi\rangle \in \mathcal{H}$ has the form (1). Let $\{|i_1 \dots i_n\rangle\}$, $i_k = 0, 1$, be the computational basis of $\mathcal{H}$. Then $|i_1 \dots i_n\rangle = \prod_{k=1}^{n} X_k^{i_k}|0\rangle$, and therefore the group is $G = N_1^{\otimes n}$, where $N_1 = \{\mathbb{1}, X\}$ is the group generated by a single spin-flip. Obviously, in this case $G \cong \mathbb{Z}_2^{\otimes n}$. The condition (10) is automatically satisfied, as $\langle 0|\prod_i X_i|0\rangle = 0$ for any product of spin-flip operators. [1]

[1] This can be generalized to an arbitrary system of qudits, since the computational basis of a single qudit Hilbert space is generated by applying the generalized Pauli x-operator on $|0\rangle$: $X|i\rangle = |i \oplus 1\rangle$, with $\oplus$ addition modulo d. The group is $G \cong \mathbb{Z}_d^{\otimes n}$.

In this case the associated G-homogeneous state is a *stabilizer state*. An n-qubit stabilizer state is invariant under the action of a group (the stabilizer) having n generators. In the following we will discuss the relationship between G-homogeneous states and stabilizer states. The connection will be made via graph states [5,6], as it has been shown that any stabilizer state is equivalent to a graph state under local Clifford operations [7].

Let $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{Z}_2^n,\ a_k = 0, 1$, be a binary n-vector. We can write an arbitrary spin flip operator on the n-qubit system as $X(\mathbf{a}) := \prod_{i=1}^n X_i^{a_i}$. The full group can be written as $G = X(\mathcal{A})$, where $\mathcal{A} \subset \mathbb{Z}_2^n$ is the group of binary vectors corresponding to all the elements of G; the group operation of $\mathcal{A}$ is addition modulo 2. We define the inner product of two binary vectors as $(\mathbf{a}, \mathbf{b}) := \sum_{i=1}^n a_i b_i \bmod 2$. Then one can prove that the stabilizer of $|G\rangle$ is [1]:

$$\mathcal{S}_{|G\rangle} = X(\mathcal{A}) \cdot Z(\mathcal{A}^\perp) \tag{11}$$

where $\mathcal{A}^\perp = \{\mathbf{b} \mid (\mathbf{a}, \mathbf{b}) = 0,\ \forall \mathbf{a} \in \mathcal{A}\}$ is the orthogonal complement of $\mathcal{A}$. Since the stabilizer generators are only of X- or Z-type, the G-homogeneous state is a Calderbank-Shor-Steane (CSS) state [8], which are locally equivalent to 2-colorable graph states [9,1]. Therefore G-homogeneous states are locally equivalent to 2-colorable graph states.

Since $\mathcal{A}$ and $\mathcal{A}^\perp$ are dual, $\dim \mathcal{A} + \dim \mathcal{A}^\perp = n$, we can choose the group G to have at most $\lfloor n/2 \rfloor$ generators ($\dim \mathcal{A}$ is the number of its generators). This can be readily seen from the following argument. Define $|G'\rangle := H^{\otimes n}|G\rangle$, where H is a Hadamard gate; obviously the two states are locally equivalent, $|G'\rangle \simeq_{LU} |G\rangle$. Since the Hadamard gate interchanges the X and Z operators, the stabilizer of the new state is $\mathcal{S}_{|G'\rangle} = X(\mathcal{A}^\perp) \cdot Z(\mathcal{A})$, hence the new spin-flip group is $G' = X(\mathcal{A}^\perp)$. Therefore, for every G-homogeneous state $|G\rangle$, with G having k generators, there is a locally equivalent G'-homogeneous state $|G'\rangle$, where G' has now $n - k$ generators. As both $|G\rangle$ and $|G'\rangle$ have the same entanglement, one can immediately derive a bound for the entropy of G-homogeneous states:

$$S(|G\rangle) \leq \min(k, n-k) \leq \lfloor n/2 \rfloor \tag{12}$$

The compact representation of 2-colorable graph states as G-homogeneous states (modulo local unitaries) enables us to easily compute a measure of n-partite entanglement, the n-tangle. The n-tangle τ_n is defined, for n even, as $\tau_n := |\langle\psi|Y^{\otimes n}|\psi^*\rangle|^2$ [10]. A simple computation for a G-homogeneous state (and hence for a 2-colorable graph state) gives for the n-tangle the following result [1]:

$$\tau_n(|G\rangle) = (1 - p(G))\ \chi_G(X^{\otimes n}) \tag{13}$$

where $p(G) = 0$ if all generators of G are even (i.e., are a product of an even number of spin-flip operators) and $p(G) = 1$ otherwise; χ_G is the characteristic function of G. Hence the n-tangle is 1 iff all the generators of G are even and $X^{\otimes n} \in G$.

References

[1] A. Hamma, R. Ionicioiu, and P. Zanardi, Phys. Rev. A **72**, 012324 (2005); quant-ph/0504049.
[2] A. Hamma, R. Ionicioiu, and P. Zanardi, Phys. Rev. A **71**, 022315 (2005); quant-ph/0409073.
[3] M. Srednicki, Phys. Rev. Lett. **71**, 666 (1993).

[4] M.B. Plenio, J. Eisert, J. Dreißig, and M. Cramer, Phys. Rev. Lett. **94**, 060503 (2005).
[5] M. Hein, J. Eisert and H.J. Briegel, Phys. Rev. A **69**, 062311 (2004).
[6] R. Raussendorf, D.E. Browne and H.J. Briegel, Phys. Rev. A **68**, 022312 (2003).
[7] M. Van den Nest, J. Dehaene, and B. De Moor, Phys. Rev. A **69**, 022316 (2004).
[8] A.R. Calderbank and P.W. Shor, Phys. Rev. A **54**, 1098 (1996); A.M. Steane, Phys. Rev. Lett. **77**, 793 (1996).
[9] K. Chen, H.K. Lo, *Conference Key Agreement and Quantum Sharing of Classical Secrets with Noisy GHZ States*, quant-ph/0404133.
[10] A. Wong and N. Christensen, Phys. Rev. A **63**, 044301 (2001); quant-ph/0010052.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

A Quantum Algorithm for Closest Pattern Matching

Paulo Mateus [a] and Yasser Omar [b,1]
[a] *Centro de Lógica e Computação, Instituto Superior Técnico*
[b] *CEMAPRE, ISEG, Universidade Técnica de Lisboa*

Abstract.
We propose a quantum algorithm for closest pattern matching which allows us to search for as many distinct patterns as we wish in a given string (database), requiring a query function per symbol of the pattern alphabet. This represents a significant practical advantage when compared to Grover's search algorithm as well as to other quantum pattern matching methods [7], which rely on building specific queries for particular patterns. Our method makes arbitrary searches on long static databases much more realistic and implementable. Our algorithm, inspired by Grover's, returns the position of the closest substring to a given pattern with high probability in $O(\sqrt{N})$ queries, where N is the size of the string. The analysis and quantum circuit construction of our algorithm are presented in [4].

Keywords. Quantum Algorithms, Database Search.

Search in databases is nowadays a common and fundamental application in computer science, one that we use daily to find a word in a text, or a site in Google. Currently we are also living the quantum information revolution, where the idea to encode information in quantum systems offers us a radically new type of information which allows for much securer communications and much faster computations than what we were able to achieve so far using (now-called) classical information [6]. There have been a few quantum algorithms proposed (the most significant probably being Shor's efficient factorization algorithm [8] of 1994, solving a problem that classically is believed to be intractable) even though the construction of a scalable quantum computer is still a challenge, presently being tackled with a plethora of different technologies [1]. Yet, should quantum computation become a reality, there is still no implementable efficient quantum algorithm to search a given database (in the current models of quantum computer), despite Grover's celebrated quantum search algorithm proposed in 1996 [3]. Grover's work, which now constitutes a paradigm for quantum search algorithms, offers a quadratic speed-up in query complexity (i.e. calls of a query function) when compared to the classical case. However, in the real execution of these search algorithms, we must distinguish the *compile time* and the *run time*. The compile time is essentially the construction of the query function on which the

[1] Correspondence to: Departamento de Matemática, Instituto Superior de Economia e Gestão, P-1200-781 Lisbon, Portugal; E-mail: yasser at qubit.org.

algorithm relies to identify the element being searched. But this construction is, in general, not a negligible task. In particular, for database search, we must go through all the database elements to build the so-called oracle, so that we can then implement the search. Note that this makes the quantum search irrelevant in practical terms, since you need to know the solution to run it. Moreover, given a query function built to find a particular element, it can only be used again to find that very same element. The search for a different item in the same database would require building a new specific query function. All this represents a serious obstacle to the application of current search algorithms [2].

To address this problem, we propose a quantum algorithm for pattern matching which allows us to search for as many distinct patterns as we wish in a given unsorted string (database), and moreover returns the position of the closest substring to a given pattern with high probability in $O(\sqrt{N})$ queries, where N is the size of the string. This means that the time to find the *closest* match (a much harder problem than to find the *exact* match, as we shall see) does not depend on the size of the pattern itself, a result with no classical equivalent. Another crucial point is that our quantum algorithm is actually useful and implementable to perform searches in (unsorted) databases. For this, we introduce a query function per symbol of the pattern alphabet, which will require a significant (though clearly efficient) pre-processing, but will allow us to perform an arbitrary amount of different searches in a static database. A *compile once, run many* approach yielding a new search algorithm that not only settles the previously existing implementation problems, but even offers the solution of a more general problem, and with a very interesting speed-up. The analysis of the algorithm as well as the recipe for its full implementation as a quantum circuit are given in [4].

In the classical setting, the best algorithm, to our knowledge, for the closest substring problem takes $\mathrm{O}(MN)$ queries where M is the size of the pattern. This result follows from adapting the best known algorithm for approximate pattern matching [5], which takes $\mathrm{O}(eN + M)$ where e is the number of allowed errors. One should not compare the closest match to (exact) pattern match, where the problem consists in determining if a certain word (pattern) is a substring of a text. For exact pattern matching it is proven that the best algorithm can achieve $\mathrm{O}(M + N)$ [5]. However, in practical cases where data can mutate over time, like DNA, or is store in a faulty systems, the closest match problem is a much more relevant, since sometimes, only approximates of the pattern exist, but nevertheless need to be found.

Our algorithm is based on the modified Grover search algorithm proposed in [2] for the case of multiple solutions. It uses the techniques originally introduced by Grover [3]: a query function that marks the state encoding the database element being searched by changing its phase; followed by an amplitude amplification of the marked state. The state can be detected with high probability by iterating this process $\sqrt{N}$ times where N is the size of the database.

[2] Yet, Grover's algorithm can be extremely useful and represent an effective speed-up for other search problems, as for instance in checking if the elements of a list are solutions of a given NP-complete problem, where the query function plays the role of the verifier and can be easily implemented.

Let us now describe our closest pattern matching algorithm. Given a string w of size N over an alphabet Σ, we want to know if a certain pattern p of size M occurs in w, or at least obtain the closest match to p in w. In particular we want to find the position $i \in \{1, \ldots, N\}$ where a certain symbol of p occurs in w. To this end, we encode position i in a unit vector $|i\rangle$ of a Hilbert space $\mathcal{H}$ of dimension N (where the set $B = \{|1\rangle, \ldots, |N\rangle\}$ constitutes an orthonormal basis of $\mathcal{H}$).

The initial state of the total system reflects the fact that we want the second symbol of p to occur just after the first, and the third to occur just after the second, and so on. For this reason we consider the following initial state, which consists of a uniform superposition of all possible states fulfilling this property: $|\psi_0\rangle = \sum_{k=1}^{N-M+1} \frac{1}{\sqrt{N-M+1}} |k\rangle$,

To perform the search, we now need to define a query function q_σ for each symbol σ of the alphabet Σ. We will thus have $|\Sigma|$ different query functions[3], each given by:

$$q_\sigma(i) = \begin{cases} 1 & \text{if the } i\text{-th letter of } w \text{ is } \sigma \\ 0 & \text{otherwise} \end{cases} . \tag{1}$$

As in Grover's algorithm, we want to use the query to mark states where there is a match for the individual symbol, in particular by shifting the phase of the respective state, as given by the following unitary transformation: $U_\sigma |k\rangle = (-1)^{q_\sigma(k)} |k\rangle$, where $|k\rangle \in B$. In our algorithm, we start by randomly choosing an element of the pattern p. Imagine we pick the j-th element of p ($j \in \{1, \ldots, M\}$). Then, whenever the $(k + j - 1)$-th letter of the string matches the j-th element of the pattern, we mark state $|k\rangle$ by shifting its phase. To produce this marking, we define the transposition operators T_j given by: $T_j |k\rangle = |k + j - 1 \bmod N\rangle$. The desired phase shifting is then obtained with $T_j^{-1} U_{p_j} T_j$.

On average, a position with a partial match, say of M' out of M matches of individual symbols, will have the phase shift applied $\frac{M'}{M}$ times. Note that the more matches we obtain, the more the phase will be shifted, and consequently the more the amplitude will be amplified. Observe that for a given string there might be full and partial matches, leading to larger and smaller amplitude amplifications respectively.

The amplitude amplification is obtained by applying the usual Grover diffusion D to the total state, over the $N - M + 1$ relevant basis:

$$D = (2(|\psi_0\rangle\langle\psi_0|) - I) + 2 \left(\sum_{k=N-M+2}^{N} |k\rangle\langle k| \right), \tag{2}$$

where I is the identity operator of dimension N.

The algorithm is then constituted by iterating the phase shift induced by the query followed by amplitude amplification. The final step is to measure the state of a symbol of the pattern over the predefined basis B, yielding the position of the closest match of the pattern in the string with a high probability.

[3]Note that, in view of implementations, it is straightforward to define the corresponding unitary operator Q_σ, acting over $\mathcal{H} \otimes \mathcal{H}_2$ (where $\mathcal{H}_2$ is the Hilbert space of dimension 2), as follows: $Q_\sigma(|i\rangle \otimes |b\rangle) = |i\rangle \otimes |q_\sigma(i) \oplus b\rangle$, where $|i\rangle$ encodes position i and $|b\rangle$ is a auxiliary qubit.

In summary, the algorithm will be as follows:

Input: $w \in \Sigma^*$ and $p \in \Sigma^*$
Output: $m \in \mathbb{N}$
Quantum variables: $|\psi\rangle \in \mathcal{H}(\{1, \ldots, N\})$
Classical variables: $r, i, j \in \mathbb{N}$

1. choose $r \in [0, \llcorner\sqrt{N - M + 1}\lrcorner]$ uniformly,
2. set $|\psi\rangle = \sum_{k=1}^{N-M+1} \frac{1}{\sqrt{N-M+1}} |k\rangle$;
3. for $i = 1$ to r
   (a) choose $j \in [1, M]$ uniformly
   (b) set $|\psi\rangle = T_j^{-1} U_{p_j} T_j |\psi\rangle$;
   (c) set $|\psi\rangle = D|\psi\rangle$
4. set m to the result of the measurement of $|\psi\rangle$ over the base $\{|1\rangle, \ldots, |N\rangle\}$.

The analysis of the algorithm is inspired by [2] and is detailed in [4], where we conclude that, in the quantum circuit model, our algorithm has an efficient compile time of $O(N \log^2(N) \times |\Sigma|)$ and a total run time of $O(N^{3/2} \log^2(N) \log(M))$.

In summary, we have presented a quantum algorithm for closest pattern matching that not only makes the quantum search in (long unsorted) static databases realistic, but even interesting, as it is a strong candidate to offer a faster solution than what is known classically for this important problem. Based on a *compile once, run many times* approach, our algorithm allows for an arbitrary amount of different searches on the same string, and offers a query complexity of $O(\sqrt{N})$ in the most relevant limit where the size M of the pattern is much smaller than the size N of the database. In [4], we give the details on how to obtain the full quantum circuit that implements our algorithm, thus offering an oracle-based quantum algorithm ready to be implemented.

The authors would like to thank A. Ambainis, L. Grover, E. Kashefi, J.I. Latorre, U. Vazirani and A. Sernadas for useful discussions, and gratefully acknowledge the support from Fundação para a Ciência e a Tecnologia (Portugal), namely through programs POCTI/POCI and project POCI/MAT/55796/2004 QuantLog, partially funded by FEDER (EU).

References

[1] ARDA, Quantum Information Science and Technology Roadmap, *qist.lanl.gov* .
[2] M. Boyer, J. Brassard, P. Høyer and A. Tapp, *Fort. Phys.*, 46(1-5):493–505, 1998.
[3] L. Grover, *Phys. Rev. Lett.*, 79:325–328, 1997.
[4] P. Mateus and Y. Omar, *arXiv.org*, quant-ph/0508237, 2005.
[5] G. Navarro, *ACM Comput. Surv.*, 33(1):31–88, 2001.
[6] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, UK, 2000.
[7] H. Ramesh and V. Vinay, *J. Discrete Algorithms*, 1(1):103–110, 2003.
[8] P. W. Shor, *SIAM J. Comp.*, 26(5):1484–1509, 1996.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Classical and Quantum Fingerprinting in the One-Way Communication Model

A. J. Scott [1], Jonathan Walgate and Barry C. Sanders
Institute for Quantum Information Science, University of Calgary, Canada

Abstract. We compare the worst-case performance of quantum fingerprinting strategies under the one-way communication model to that of classical strategies when the error is one-sided. Bounds on the classical worst-case error probability are reported, while quantum strategies surpassing these bounds are constructed from spherical codes, equiangular tight frames and mutually unbiased bases.

Keywords. quantum fingerprinting, communication complexity

Consider the following scenario. Alice needs to send an important message to Bob. She does not mind if an intermediary party reads this message while in transit, but does however, require that the message remains unchanged. Alice considers broadcasting her message through a trusted public communication channel, her local radio station, but alas, access time is very expensive and her message is long. Instead she proceeds by sending it across untrusted communication channels, over the internet, which are inexpensive and plentiful. Alice is now faced, however, with the possibility that a saboteur has surreptitiously intercepted and changed part, or all, of her message. To test its integrity, she reluctantly reconsiders buying airtime at her local radio station. Alice's transmission costs now depend on the level of risk she is prepared to tolerate, for a corrupted message to pass undetected. An absolute guarantee of integrity requires Alice to transmit her entire message again, this time through the trusted channel. If however, she is willing to tolerate a small probability of detection error, a considerably smaller piece of information, called a fingerprint, need only be sent. Alice could, for example, broadcast only the first byte of her message. This strategy is said to have one-sided error: when the first byte does not match that for Bob's copy, Alice knows with certainty that foul play has occurred. Unfortunately, since it is Bob who must perform the comparison, and no private communication channel is available, we are forced to assume that Alice and Bob use a strategy which is publicly known. An adversary will now simply avoid corrupting the first byte, while playing havoc with the rest. There is, of course, a simple workaround: Alice instead broadcasts a random byte of her message, chosen by flipping a private coin, together with the location of this byte. Assuming the availability of a trusted, though costly, broadcast channel, Alice has thus obtained a level of

[1]Correspondence to: Andrew Scott, Department of Physics and Astronomy, University of Calgary, Calgary AB T2N1N4, Canada. E-mail: ascott@qis.ucalgary.ca

Table 1. The minimum achievable worst-case error probability under the one-way communication model for a classical fingerprinting strategy which maps n messages to m fingerprints with one-sided error.

$n\backslash m$	2	3	4	5	6	7	8	9	10
2	0	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0	0
4	1	1	0	0	0	0	0	0	0
5	1	1	1/2	0	0	0	0	0	0
6	1	1	1/2	1/2	0	0	0	0	0
7	1	1	1	1/2	1/2	0	0	0	0
8	1	1	1	1/2	1/2	1/2	0	0	0
9	1	1	1	1/2	1/2	1/2	1/2	0	0
10	1	1	1	1/2	1/2	1/2	1/2	1/3	0
11	1	1	1	1	1/2	1/2	1/2	1/3	1/3
12	1	1	1	1	1/2	1/2	1/2	1/3	1/3
13	1	1	1	1	1/2	1/2	1/2	1/2	1/3
14	1	1	1	1	1/2	1/2	1/2	1/2	1/2
15	1	1	1	1	1/2	1/2	1/2	1/2	1/2
16	1	1	1	1	2/3	1/2	1/2	1/2	1/2

security through only public communication. In a probabilistic sense, message sabotage will never go undetected.

Although motivated in terms of message authentication, the preceding protocol is called *fingerprinting*. In all generality, fingerprinting enables separated parties to infer whether the messages they hold are the same or different when the cost of communication is high: each message is associated with a smaller fingerprint, and comparisons between messages are made in terms of their fingerprints alone. In the above authentication scheme, the messages were compared using one-way communication, from Alice to Bob. It is known from the theory of communication complexity [1], that in the limit of long messages, it is sufficient, and necessary, for Alice to communicate fingerprints of length $O(\log N)$ bits, for messages of length N bits, if a small constant error probability is allowed. This error probability can then be made arbitrarily small through repeated use of the protocol. In another model of communication complexity, called the simultaneous message passing model [2], the fingerprints are transmitted by both Alice and Bob, but this time to a referee, who decides the outcome of the message comparison. In this case it is sufficient, and necessary, for Alice and Bob to communicate fingerprints of length $O(\sqrt{N})$ bits if the error is to be kept arbitrarily small [3].

Quantum fingerprinting protocols allow each classical fingerprint to be replaced by a quantum state. For a fair comparison, the quantum fingerprints are drawn from a Hilbert space of dimension equal to the number of available classical fingerprints. In the simultaneous message passing model it is known that fingerprints composed of quantum information can be made exponentially smaller than those composed of classical information. Specifically, for messages of length N bits, it was shown by Buhrman *et al.* [4] that fingerprints composed of $O(\log N)$ qubits are sufficient to keep the error arbitrarily small. This exponential resource advantage is not apparent in the one-way communication model. Nonetheless, in this partial summary of Ref. [5], we compare the worst-case performance of quantum fingerprinting strategies for short messages, under the one-way communication model, to that of classical strategies, when the error is one-sided.

Classical one-way fingerprinting can be described by the following scenario. A supplier, called Sapna, chooses two messages, x and y, from a pool of n unique

Table 2. The smallest known maximal pairwise overlap for sets of n pure states in $\mathbb{C}^m$. Each entry is also an achievable worst-case error probability under the one-way communication model for a quantum fingerprinting strategy which maps n messages to m fingerprints with one-sided error.

$n\backslash m$	2	3	4	5	6	7	8	9	10
2	0	0	0	0	0	0	0	0	0
3	$\mathbf{1/4}^e$	0	0	0	0	0	0	0	0
4	$\mathbf{1/3}^e$	$\mathbf{1/9}^e$	0	0	0	0	0	0	0
5	**1/2**	·1886	$\mathbf{1/16}^e$	0	0	0	0	0	0
6	$\mathbf{1/2}^m$	$\mathbf{1/5}^e$	·1071	$\mathbf{1/25}^e$	0	0	0	0	0
7	**·6051**	$\mathbf{2/9}^e$	$\mathbf{1/8}^e$	·0709	$\mathbf{1/36}^e$	0	0	0	0
8	**·6306**	1/4	$\mathbf{1/7}^e$	·0871	·0502	$\mathbf{1/49}^e$	0	0	0
9	**2/3**	$\mathbf{1/4}^e$	·1615	·1025	$\mathbf{1/16}^e$	·0384	$\mathbf{1/64}^e$	0	0
10	**·7022**	**1/3**	·1687	$\mathbf{1/9}^e$	·0741	·0491	·0299	$\mathbf{1/81}^e$	0
11	**·7236**	**1/3**	·1808	$\mathbf{3/25}^e$	$\mathbf{1/12}^e$	·0573	·0394	·0238	$\mathbf{1/100}^e$
12	**·7236**	$\mathbf{1/3}^m$	·1830	·1277	$\mathbf{1/11}^e$	·0651	·0456	·0319	·0193
13	**·7713**	·3871	$\mathbf{3/16}^e$	·1351	·0972	·0714	·0522	$\mathbf{1/27}^e$	·0265
14	**·7820**	·4066	1/5	·1411	·1026	$\mathbf{1/13}^e$	·0577	·0429	·0310
15	·7963	·4141	1/5	·1452	·1072	$\mathbf{4/49}^e$	$\mathbf{1/16}^e$	·0476	·0357
16	·8061	·4196	$\mathbf{1/5}^e$	·1506	$\mathbf{1/9}^e$	·0857	$\mathbf{1/15}^e$	·0519	$\mathbf{1/25}^e$

messages and hands them to Alice and Bob, respectively. Alice and Bob are now tasked with determining whether the messages they hold are the same or different using one-way communication only, from Alice to Bob say. As communication is considered expensive, Alice is limited to sending a fingerprint of her original message, a, which she selects from a smaller pool of size m. Bob then infers

$$\mathrm{EQ}(x,y) = \begin{cases} 1, \text{if } x = y \\ 0, \text{if } x \neq y \end{cases}$$

and completes the protocol by announcing a single bit $z \in \{0,1\}$. Bob is correct if $z = \mathrm{EQ}(x,y)$. In the current investigation we are concerned only with one-sided-error protocols, in which case, $z = 0$ is allowed only if $x \neq y$. Protocols with one-sided error are of vital practical importance whenever the cost of false negative results outweighs that of false positives. This is the case when x and y are in fact expected to be equal. In the authentication scheme outlined above, the one-sided-error condition means that Bob will always accept an uncorrupted message as authentic, and thus, transmission throughput remains unaffected when an adversary is not present.

The fingerprinting protocol adopted by Alice and Bob is publicly announced. The goal of this protocol is to minimize Bob's error probability. Sapna, however, may be a saboteur, and always choose message pairs that lead to the highest error rate in Bob's output. We thus evaluate fingerprinting protocols according to this worst-case scenario. The worst-case error probability, $P_{\mathrm{wce}} = \max_{x,y} \mathrm{Prob}(z \neq \mathrm{EQ}(x,y))$, then corresponds to the maximum error rate, and provides an absolute guarantee on the protocol's performance. For small m and n, the minimum achievable worst-case error probability for a classical one-way fingerprinting strategy with one-sided error is reported in Table 1 [5].

In the quantum fingerprinting scenario, Alice's classical fingerprints are replaced by m-dimensional mixed quantum states, $\rho(x)$. Quantum fingerprinting becomes classical when $\rho(x)$ is diagonal in the computational basis $\{|a\rangle\}$, in which case we may identify $p(a|x) = \langle a|\rho(x)|a\rangle$ as the probability that Alice sends fingerprint a to Bob, given that she receives message x from Sapna. Upon receipt of Alice's fingerprint, Bob's task is to determine whether $\rho(x)$ is $\rho(y)$ or not, with

one-sided error, and thus the projective measurement $\{P(y), 1 - P(y)\}$, where $P(y)$ projects onto the support of $\rho(y)$, is the only suitable choice.

It is intuitively reasonable that Alice should fingerprint with pure states, and thus, we will restrict our study to this case. The worst-case error probability is then simply the maximal pairwise overlap for Alice's choice of fingerprint states $\{|\psi(x)\rangle\}_{x=1}^{n} \subset \mathbb{C}^m$: $P_{\text{wce}} = \max_{x \neq y} |\langle\psi(x)|\psi(y)\rangle|^2$. Consequently, the minimum possible worst-case error probability will be achieved when Alice chooses her set of fingerprint states to have the minimum possible maximal pairwise overlap. In Table 2 we report the smallest known maximal pairwise overlap for sets of n pure states in $\mathbb{C}^m$. Entries in boldface signify that the overlap is optimal. This is the case for all *equiangular tight frames* (ETFs) [6], denoted by a superscript e in the table, and *mutually unbiased bases* (MUBs) [7], denoted by an m. When $m = 2$, through the Bloch-sphere representation of a qubit, our problem is equivalent to finding optimal *spherical codes* [8] in three dimensions. Tables of putatively optimal spherical codes [9] give the maximal overlaps in the first column of Table 2. Of these, only in the cases $n \leq 14$ has optimality been established. All other entries in Table 2 are the overlaps for the best packings found in a numerical search, and thus, optimality is not guaranteed. The value given, however, is likely to be optimal, or at least very close to optimal. In fact, in many of the cases where $n \leq m^2$, the given overlap almost saturates a known lower bound [5].

Each of the maximal overlaps in Table 2 is the worst-case error probability for a quantum fingerprinting strategy. Comparing these to the classical error probabilities in Table 1, we see that quantum strategies generally outperform classical strategies by a sizeable margin. For example, when $n = 6$ the optimal classical strategy requires $m = 4$ fingerprints to obtain $P_{\text{wce}} = 1/2$. This error probability may be reduced to $P_{\text{wce}} \approx 0 \cdot 1071$, however, using quantum fingerprints. Taking a different approach, if we instead fix $n = 6$ and $P_{\text{wce}} = 1/2$, a quantum strategy exists which achieves this performance using only $m = 2$ fingerprints. Finally, the quantum performance gain is more impressive if we instead fix m and P_{wce}. In fact, a numerical search finds that there exists a set of n 4-dimensional quantum states with a maximum pairwise overlap of $1/2$, or less, for all $n \leq 75$. Thus, quantum fingerprinting strategies can map at least 75 messages into 4 fingerprints while retaining $P_{\text{wce}} \leq 1/2$, compared to only 6 messages classically.

References

[1] E. Kushilevitz and N. Nisan, *Communication Complexity* (Cam. Univ. Press, 1997).
[2] A. C.-C. Yao, in *Proc. 11th ACM Symp. Theory Comp.* (ACM, 1979).
[3] A. Ambainis, Algorithmica **16**, 298 (1996); I. Newman and M. Szegedy, in *Proc. 28th ACM Symp. Theory Comp.*, (ACM, 1996).
[4] H. Buhrman *et al.*, Phys. Rev. Lett. **87**, 167902 (2001).
[5] A. J. Scott, J. Walgate and B. C. Sanders, e-print quant-ph/0507048.
[6] T. Strohmer and R. W. Heath Jr., Appl. Comp. Harm. Anal. **14**, 257 (2003).
[7] W. K. Wootters and B. D. Fields, Ann. Phys. **191**, 363 (1989).
[8] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups* (Springer, 1988).
[9] N. J. A. Sloane *et al.*, *Spherical Codes*, http://www.research.att.com/~njas/packings/

This page intentionally left blank

Chapter 3

Quantum Information Theory in Spin Systems

This page intentionally left blank

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Introduction to Localizable Entanglement

Markus Popp [a,1], Frank Verstraete [b], Miguel Angel Martin-Delgado [c] and Ignacio Cirac [a]

[a] *Max-Planck-Institut für Quantenoptik, 85748 Garching, Germany*
[b] *Institute for Quantum Information, Caltech, Pasadena, CA 91125 USA*
[c] *Departamento de Física Teórica I, Universidad Complutense de Madrid, E-28040, Spain*

Abstract. We present here an introduction to the concept of localizable entanglement (LE). It is defined as the maximum entanglement that can be localized, on average, between two parties of a multipartite system by performing local measurements on the remaining parties. This entanglement measure leads naturally to notions like entanglement length and entanglement fluctuations. For both spin-1/2 and spin-1 systems we prove that the LE of a pure quantum state can be lower bounded by connected correlation functions. We further propose a scheme, based on matrix-product states and the Monte Carlo method, to efficiently calculate the LE for quantum states of a large number of spins. The virtues of LE are illustrated for various spin models. In particular, characteristic features of a quantum phase transition such as a diverging entanglement length can be observed. We also give examples for pure quantum states exhibiting a diverging entanglement length but finite correlation length. We have numerical evidence that the ground state of the antiferromagnetic spin-1 Heisenberg chain can serve as a perfect quantum channel. Furthermore, we apply the numerical method to mixed states and study the entanglement as a function of temperature.

1. Introduction

The creation and distribution of entangled states plays a central role in quantum information, because it is the key ingredient for performing certain quantum information tasks, like teleportation [1] or quantum computation. In this respect, multiparticle quantum states can be considered as entanglement resources, naturally appearing in many physical systems. On the other hand, it is believed that the study of multipartite entanglement might prove fruitful in other fields of physics, like condensed matter, e.g. for understanding the complex physics of strongly correlated states [2,3]. In particular it has been shown [4,5,6,7,8,9] that the ground state entanglement of various spin systems may exhibit characteris-

[1] Correspondence to: Markus Popp, MPQ, Hans-Kopfermann-Str. 1, D-85748 Garching, Germany. Tel.: +49 (0) 89 32905-368; Fax: +49 (0) 89 32905-336; E-mail: markus.popp@mpq.mpg.de.

tic features at a quantum phase transition. Hence it is desirable to find ways of characterizing and quantifying entanglement in multipartite systems.

In [10,11] an entanglement measure, called the *Localizable Entanglement* (LE), has been defined. It quantifies the bipartite entanglement contained in a multipartite system. The concept of LE allows one to define the notion of entanglement length, which characterizes the typical distance up to which bipartite entanglement can be localized in the system. Moreover, it has been shown for general pure qubit and qutrit states that the LE can always be lower bounded by connected correlation functions [10,11]. Hence, quantum phase transitions, characterized by a diverging correlation length, are equivalently detected by a diverging entanglement length. In fact, the concept of LE has already proven useful in several studies of entanglement properties of spin systems [10,11,12,13,14,15]. Indeed, for ground states of many spin-1/2 models one finds that correlations and entanglement (as given by LE) typically exhibit the same qualitative behavior. In gapped spin-1 systems, which exhibit finite correlation lengths [16], however, a ground state has been found for which the entanglement length is infinite [12]. This example also shows that the concept of LE can serve to detect hidden order in certain states.

The LE is the maximum entanglement that can be localized, on average, between two parties of a multipartite system, by performing local measurements on the other parties. Hence, LE is defined in an operational way and has a clear physical meaning. For instance it can be used as a figure of merit to characterize the performance of quantum repeaters [17]. Note that, in the context of LE, particles are not traced out but measured. This is in contrast to earlier approaches, where the concurrence of the reduced density operator of two separated spins in a spin chain has been calculated (see e.g. [4,5]). Although the concurrence exhibited characteristic features at a quantum phase transition, it does not detect long range quantum correlations.

The fundamental difference between tracing and measuring can be illustrated with two simple examples: For both the GHZ state [18] and the cluster state [19] it can readily be checked that the reduced density operator of any two qubits contains no entanglement at all. On the other hand one can find a local measurement basis, such that the LE is maximal. We further note that in the case of the GHZ-state the entanglement properties could also have been revealed by studying the connected version of the two-point correlation function. Indeed, the GHZ-state is one of many examples [10], for which correlations can be identified with quantum correlations, i.e. entanglement. However, this intimate connection does not hold true for all pure quantum states (as shown in [12]). For instance, in the case of the cluster state the LE is maximal, whereas the connected correlations are all zero.

The intention of this paper is to provide an overview of the concept of LE as an measure for localizable bipartite entanglement of multipartite quantum systems [10,11]. Starting from the original definition, we introduce several variants of LE, and establish basic relations between them. Apart from the notion of entanglement length , we also give a meaning to the notion of *entanglement fluctuations* in terms of LE. We prove for the case of pure qubit and qutrit states that LE can be lower bounded by connected two-particle correlation functions. Apart from these analytical findings we present a method for the numerical com-

putation of the LE for arbitrary one dimensional spin systems. It allows us to efficiently simulate chains with even more than 100 sites and also works for finite temperatures. We apply this method for the study of LE present in ground states of various spins models. We find that the entanglement fluctuations as well as the LE exhibit characteristic features at a quantum phase transition. For instance, we observe a discontinuity in the first derivative of LE at a Kosterlitz-Thouless transition [20]. These findings can be understood as a direct consequence of the numerical observation, that for spin-1/2 systems the LE and the maximal connected correlation function are typically equal to each other. In the case of the spin-1 Heisenberg antiferromagnet we observe rather the opposite behavior, namely an infinite entanglement length but finite correlation length. In this context we also comment on a possible connection between the presence of hidden order in the state and the existence of long range entanglement. In particular, we present examples showing that, in general, such a connection does not exist. As a further application of our numerical method we study the effect of finite temperature on the LE. For the AKLT model [21] we find that the entanglement length increases exponentially with the inverse temperature.

The paper is organized as follows: We begin with reviewing algebraically computable entanglement measures for bipartite systems. Then we give a formal definition of LE and derive quantities like entanglement length and entanglement fluctuations. We further connect the concept of LE with the idea of quantum repeaters. In Sect. 4 we derive a lower bound of LE in terms of connected correlation functions. Next, we present in Sect. 5 a numerical method to compute the LE for spin chains. The numerical scheme is based on the matrix-product state (MPS) [22,23] representation of ground states and the Monte Carlo (MC) [24] method. In Sect. 6 this method is applied to the calculation of LE and the entanglement fluctuations for ground states of various standard spin-1/2 chains. In Sect. 7 we study the LE of gapped spin-1 models. In particular, we present numerical calculations for the ground state of the Heisenberg antiferromagnet, showing that the entanglement length diverges. Furthermore, we comment on a possible connection between a diverging entanglement length and hidden order in the system. Finally, we demonstrate that our numerical method can also be used to calculate the LE of mixed states. As an example we compute the entanglement length as a function of the temperature for the AKLT model.

2. Entanglement of Bipartite Systems

In this section we briefly review the entanglement of bipartite systems of finite dimension. Physical states of such systems are described by density matrices ρ ($\rho = \rho^\dagger, \rho > 0, \mathrm{tr}\rho = 1$) acting in the Hilbert space of the composite system $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. A state ρ is defined as *separable* if and only if it can be decomposed as a convex combination of product states:

$$\rho = \sum_i p_i \ \rho_i^A \otimes \rho_i^B, \tag{1}$$

with $\sum_i p_i = 1$ and $p_i > 0$. Inversely, a state ρ is called *entangled* if and only if such a decomposition does not exist. Often it is also desirable to quantify the *degree* of entanglement. For pure states ($\mathrm{tr}\rho^2 = 1$) a widely accepted entanglement measure exists, which can, in principle, be calculated for any Hilbert space dimension. In the case of mixed states ($\mathrm{tr}\rho^2 < 1$), however, closed expressions for the entanglement of a state ρ can typically be found only for low dimensional systems. In the following we give a summery of algebraically computable entanglement measures, which are used in this article.

2.1. Pure States

Any pure state of a bipartite system can be written in the Schmidt decomposition:

$$|\psi\rangle = \sum_{i=1}^{n} \lambda_i \, |\phi_i^A\rangle \otimes |\phi_i^B\rangle, \qquad (\lambda_i > 0) \tag{2}$$

where $\{|\phi_i^A\rangle\}$ and $\{|\phi_i^B\rangle\}$ are sets of orthonormal states for subsystems A and B. The Schmidt coefficients λ_i represent the features of the state $|\psi\rangle$, which are invariant under local unitary operations on the subsystems. By definition, the entanglement of a state cannot be increased by local operations and classical communication (LOCC). This characteristic distinguishes quantum correlations from classical correlations. As a consequence, any entanglement measure must be invariant under local unitary operations and thus should depend solely on the Schmidt coefficients λ_i. In this context it is useful to consider the reduced density operators $\rho_A = \mathrm{tr}_B|\psi\rangle\langle\psi|$ and $\rho_B = \mathrm{tr}_A|\psi\rangle\langle\psi|$ of subsystem A and subsystem B, respectively. The degree of mixing of ρ_A and ρ_B is directly related to the entanglement of $|\psi\rangle$. Hence, the *entropy of entanglement* [26] has been adopted as the standard measure of entanglement for pure states. It is defined as:

$$E_E(\psi) := S(\rho_A) = S(\rho_B) = -\sum_{i=1}^{n} \lambda_i^2 \log_2 \lambda_i^2. \tag{3}$$

Here, $S(\rho) = -\mathrm{tr}(\rho \log_2 \rho)$ denotes the von-Neumann entropy. This entanglement measure has been inspired by the observation that entanglement decreases our knowledge of local properties of the state. Another justification for this measure arises from the fact that a maximally entangled state and LOCC can be used to produce an arbitrary state $|\phi\rangle$. In this sense, the entropy of entanglement of $|\phi\rangle$ measures the "entanglement cost" of creating this state. To be more precise, one needs about $n \times E_E(|\phi\rangle)$ copies of the singlet state, $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$, for producing n copies of an arbitrary two qubit state $|\phi\rangle$.

The definition of entropy of entanglement holds for any dimension. In the special case of a 2×2 system an alternative entanglement measure is widely used, namely the *concurrence* [27]. For a pure state $|\psi\rangle$ it is defined as

$$C(\psi) = |\langle\psi^*|\sigma_y \otimes \sigma_y|\psi\rangle|, \tag{4}$$

where the complex conjugation is performed in the standard basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

For qubits it can be shown that the entropy of entanglement is a convex, monotonously increasing function, $E_E = f(C)$, of the concurrence, with

$$f(C) := H\left(\frac{1+\sqrt{1-C^2}}{2}\right), \tag{5}$$

$$H(x) := -x\log_2 x - (1-x)\log_2(1-x), \tag{6}$$

and $H(x)$ being the Shannon entropy. In this article we will typically use the concurrence to measure the entanglement of a pure two qubit state $|\,\psi\rangle$, because it can most easily be related to the maximum connected correlation function of this state [11].

2.2. Mixed States

Any mixed state can be expressed as a convex sum over pure states:

$$\rho = \sum_i p_i \;|\,\psi_i\rangle\langle\psi_i\,|, \tag{7}$$

with $p_i > 0$ and $\sum_i p_i = 1$. Unfortunately, this pure state decomposition is not unique. This means that the entanglement of ρ is not simply given by the weighted sum over the pure state entanglement. Instead, one has to take the infimum with respect to all pure state decompositions. Using this convex roof construction the *entanglement of formation* [28] of a mixed state ρ is defined as:

$$F(\rho) = \inf_{\{p_i,\psi_i\}} \sum_i \; p_i \; E_E(\psi_i). \tag{8}$$

In higher dimensions the underlying optimization problem is unsolved for general states. Only for 2×2 systems an algebraic solution exists. In the convex roof (8) one replaces E_E by the concurrence C (see Eq. (4)). This leads to the closed expression [27]:

$$C(\rho) = \max\{0, r_1 - r_2 - r_3 - r_4\}, \tag{9}$$

where the $r_i's$ are the square roots of the eigenvalues of the matrix $\rho\tilde{\rho}$ in descending order. Here the spin flipped matrix $\tilde{\rho}$ is defined as:

$$\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y), \tag{10}$$

with the complex conjugation again taken in the standard basis. Making use of Eq. (5), it has been shown that the entanglement of formation in 2×2 systems can be expressed by the concurrence: $F(\rho) = f(C(\rho))$ [27].

For higher dimensional systems one can use the *negativity* [29] as entanglement measure. Its definition relies on the $NPPT$ criterion for entanglement [30,31]: if the partial transpose ρ^{T_A} has at least one negative eigenvalue, then the state ρ is entangled. If we express ρ in an orthonormal product basis,

$$\rho = \sum_{i,j} \sum_{k,l} \langle i,k\,|\rho|\,j,l\rangle \;|\,i\rangle_A\langle j\,|\otimes|\,k\rangle_B\langle l\,|, \tag{11}$$

then, the partial transpose with respect to system A is given by:

$$\rho^{T_A} = \sum_{i,j} \sum_{k,l} \langle i,k\,|\rho|\,j,l\rangle \;|\,j\rangle_A\langle i\,|\otimes|\,k\rangle_B\langle l\,|. \tag{12}$$

We define the negativity $\mathcal{N}$ of a state ρ as

$$\mathcal{N}(\rho) = ||\rho^{T_A}||_1 - 1, \tag{13}$$

where $||.||_1$ is the 1-norm, given by the sum of the moduli of the eigenvalues. Since transposition is trace preserving, $\mathcal{N}(\rho)$ is given by two times the sum of the moduli of the negative eigenvalues of ρ^{T_A}. With this definition, the negativity of the singlet state is one. The negativity has the big advantage that it can be calculated algebraically in any dimension. However, for systems larger than 2×3 it vanishes for some entangled states [32,33] and thus it is not an entanglement measure in the strict sense. But since it is non-increasing under LOCC , it is an entanglement monotone [29].

3. Localizable Entanglement of Multipartite Systems: Definition and Basic Properties

We consider a multipartite system composed of N particles. With each particle we associate a finite dimensional Hilbert space. For simplicity we refer in the following to the particles as spins.

3.1. Definition

The Localizable Entanglement (LE) of a multi-spin state ρ is defined as the maximal amount of entanglement that can be created (i.e. localized), on average, between two spins at positions i and j by performing local measurements on the other spins. More specifically, every measurement $\mathcal{M}$ specifies a state ensemble $\mathcal{E}_{\mathcal{M}} := \{p_s, \rho_s^{ij}\}$. Here p_s denotes the probability to obtain the (normalized) two-spin state ρ_s^{ij} for the outcome $\{s\}$ of the measurements on the $N-2$ remaining spins. The average entanglement for a specific $\mathcal{M}$ is then given by:

$$\overline{L}_{i,j}^{\mathcal{M},E}(\rho) := \sum_s p_s \; E(\rho_s^{ij}), \tag{14}$$

where $E(\rho_s^{ij})$ is the entanglement of ρ_s^{ij}. Suitable entanglement measures will be discussed later in this section. The Localizable Entanglement is defined as the largest possible average entanglement:

$$L_{i,j}^{\mathcal{C},E}(\rho) := \sup_{\mathcal{M}\in\mathcal{C}} \sum_s p_s \; E(\rho_s^{ij}), \tag{15}$$

with $\mathcal{C}$ denoting the class of allowed measurements. We call the measurement $\mathcal{M}$ which maximizes the average entanglement the *optimal basis*. It is important to note that the only restriction on $\mathcal{M}$ is that the measurements are performed locally i.e. on individual spins. Apart from that, the measurment basis is arbitrary and can also vary from site to site. We distinguish three classes $\mathcal{C}$ of measurements: projective von-Neumann measurements (PM), those corresponding to positive operator-valued measures (POVM), and general local measurements that allow also for classical communication of measurement results (LOCC). In terms of LE the following relationship between these classes holds [25]:

$$L_{i,j}^{\mathrm{PM},E}(\rho) \leq L_{i,j}^{\mathrm{POVM},E}(\rho) \leq L_{i,j}^{\mathrm{LOCC},E}(\rho). \tag{16}$$

In this paper we will be mainly concerned with projective measurements. To simplify the notation we omit in this case the superscript PM.

The definition (15) still leaves open the choice of the entanglement measure E for the states $\{\rho_s^{ij}\}$. In the previous section we have listed algebraically computable entanglement measures for small dimensional systems.

In the case of a *pure* multipartite state ρ, the states $\{\rho_s^{ij}\}$ after the measurements are also pure. Hence, in any dimension we can use the entropy of entanglement E_E (3) as entanglement measure. For pure qubit states we typically choose the concurrence C (4). Due to the convexity of the function f (5) one can find the following relation for the LE as measured by these two quantities:

$$f\left(L_{ij}^{C}\right) \leq L_{ij}^{E_E} \leq L_{ij}^{C}. \tag{17}$$

Therefore the qualitative behavior of these two variants of LE will be very similar. In the case of *mixed* qubit states we make use of the closed expression for the concurrence (9). For higher dimensional mixed states we refer to the negativity $\mathcal{N}$ (13).

3.2. Entanglement Length and Fluctuations

In the field of strongly correlated systems and more specifically in the study of quantum phase transitions, the correlation length, ξ_C, is of great importance. The concept of LE readily lends itself to define the related *entanglement length*, ξ_E, as the typical length scale at which it is possible to create Bell states by doing local measurements on the other spins:

$$\xi_E^{-1} := \lim_{n\to\infty} \left(\frac{-\ln L_{i,i+n}^{E}}{n} \right). \tag{18}$$

The entanglement length is finite iff the LE $L_{i,i+n}^{E} \to \exp(-n/\xi_E)$ for $n \to \infty$, and the entanglement length ξ_E is defined as the constant in the exponent in the limit of an infinite system (see also Aharonov [34]).

Let us now have a closer look at the statistical nature of LE, as it is defined as an average over all possible measurement outcomes (see (15)). For practical purposes one can only control the measurement basis but not a specific outcome.

Therefore it would be useful to have an estimate of how much the entanglement of a particular measurement outcome deviates from the mean value as given by the LE. This information is contained in the variance of the entanglement remaining after measurements. We can thus define the notion of *entanglement fluctuations*:

$$(\delta L_{i,j}^{\mathcal{M},E})^2 := \left(\sum_s p_s \; E(\rho_s^{ij})^2 \right) - L_{i,j}^{\mathcal{M},E^2}. \tag{19}$$

The entanglement fluctuations can be defined for any measurement $\mathcal{M}$. Typically we choose for $\mathcal{M}$ the optimal basis, which maximizes the average entanglement. In this case we drop the index $\mathcal{M}$ in (19).

The study of both the entanglement length and the entanglement fluctuations could provide further inside in the complex physics of quantum phase transitions by revealing characteristic features at the quantum critical point. Examples for this are presented in Sect. 6.

3.3. Connection to Quantum Repeaters

So far we have given a purely mathematical definition of LE (15). However, it is evident that the LE is defined in an operational way that can directly be implemented on certain physical systems. In addition the concept of LE may also play an essential role in some interesting applications of quantum information theory. To be more precise, LE can serve as a figure of merit for the "performance" of certain kinds of quantum repeaters (QR).

Many tasks in quantum information processing require long-distance quantum communication. This means quantum states have to be transmitted with high communication fidelity via a quantum channel between two distant parties, Alice and Bob. Since quantum transport is also possible via teleportation [1] this problem is equivalent to establishing nearly perfect entanglement between two distant nodes. All realistic schemes for quantum communication are presently based on the use of photonic channels. However, the degree of entanglement generated between distant sites typically decreases exponentially with the length of the connecting physical channel, due to light absorption and other channel noise. To overcome this obstacle the concept of quantum repeaters has been introduced [17]. The central idea is to divide the channel into segments and to include additional nodes. Entanglement between adjacent nodes can be extended to larger distances using entanglement swapping followed by purification. After several rounds one obtains a pair of almost maximally entangled nodes, shared by Alice and Bob, that can be used for perfect quantum transport via teleportation. A possible physical realization of the QR using trapped atoms is sketched in Fig. 1 [35].

Let us now discuss how a QR setup can be characterized by the LE. First of all, let us identify the particles sitting at different nodes by spins. It is important to note that, by combining several spins to a larger Hilbert space of dimension d, the operations required for purification and entanglement swapping on this set of spins can be interpreted as local operations on a single spin of dimension d. Thus the QR can be treated as a system of interacting spins being in a state ρ. In order to assess and quantify the usefulness of such a setup as a QR one has to compute

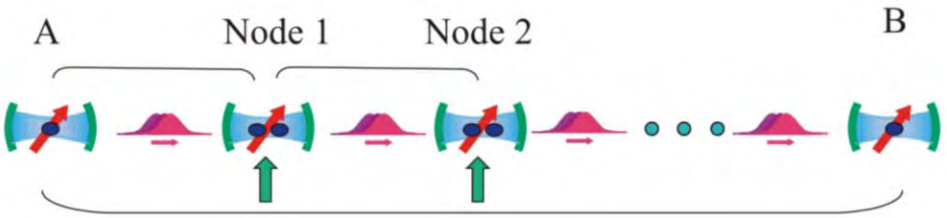


Figure 1. Illustration of the quantum repeater scheme for trapped atoms connected by optical fibres. Qubits are represented by the internal states of the atoms. Applying laser beams the internal states of atoms in adjacent cavities become entangled via the transmission of photonic states. Collective measurements on the nodes (indicated by arrows) followed by purification lead to the generation of a nearly perfectly entangled pair of qubits between A and B.

the following figure of merit: What is the maximum amount of entanglement that can be generated between the two end spins by performing local operations on the intermediate spins? But this number is nothing else than the LE. The question, which variant of LE (16) should be used, depends on the class of available local operations (PM, POVM or LOCC). Typically classical communication is allowed so that L_{ij}^{LOCC} has to be taken as figure of merit. However, not every measurement might be physically realizable. Therefore the LE will in general give an upper bound for the performance of a given QR setup.

4. Bounds on LE

Due to its variational definition, the LE is very difficult to calculate in general. Moreover, typically one does not have an explicit parameterization of the state under interest, but just information about the classical one- and two-particle correlation functions (which allows one to parameterize completely the two-spin reduced density operator ρ_{ij}). It would therefore be interesting to derive tight upper and lower bounds to the LE solely based on this information.

4.1. Upper Bound

The upper bound can readily be obtained using the concept of entanglement of assistance (EoA) [36]. The EoA depends only on the reduced density operator ρ_{ij} for spins i and j and can be defined for any spin dimension. In the case of a pair of qubits an explicit formula for EoA can be derived [37]. Given ρ_{ij} and a square root X, $\rho_{ij} = XX^\dagger$, then the EoA as measured by the concurrence reads:

$$E_A^{i,j}(\rho_{ij}) := \mathrm{tr}|X^T(\sigma_y^i \otimes \sigma_y^j)X| \ , \tag{20}$$

with $|A| = \sqrt{A^\dagger A}$. Hence $E_A^{i,j}(\rho_{ij}) = \sum_{k=1}^4 \sigma_k$, where σ_k are the singular values of the matrix $X^T(\sigma_y^i \otimes \sigma_y^j)X$. Note that for pure four qubit states a variant of EoA with local measurements was considered in [37].

4.2. Lower Bound

First, from the definition of LE (15) it follows that any specific measurement, e.g. in the computational basis, trivially provides a lower bound on the LE. More interestingly, it has been proven in [10] that for general *pure* qubit states the LE can be lower bounded by connected correlation functions. In the following we recapitulate the central results leading to this bound for spin-1/2 systems and present some extensions. Next we show to what extent these findings can be generalized to higher dimensional spin systems.

4.2.1. Spin-1/2 Systems

The basic idea is to establish a connection between the LE and connected correlation functions of the form:

$$Q^{ij}_{AB} = \mathrm{tr}[\rho(\ S^i_A \otimes S^j_B)] - \mathrm{tr}[\rho(\ S^i_A \otimes \mathbb{1})]\mathrm{tr}[\rho(\mathbb{1} \otimes S^j_B)]. \tag{21}$$

For qubits the operators S_A, S_B can be parameterized by directions $\vec{a}, \vec{b}$, representing unit vectors in a 3D real space: $S_A = \vec{a}\cdot\vec{\sigma}, S_B = \vec{b}\cdot\vec{\sigma}$ with $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. We start out by quoting the central result of [10]:
(1.i) Given a (pure or mixed) state of N qubits with connected correlation function Q^{ij}_{AB} between the spins i and j and directions $\vec{a}, \vec{b}$, then there always exists a basis in which one can locally measure the other spins such that this correlation does not decrease, on average.
An extended version of the proof is presented in [11]. There we also show that this result can be generalized to a setup, where the spins i and j can be of any dimension, but the remaining spins (on which the measurements are performed) are still qubits. In a spin-1/2 system such a situation can arise, for example, when considering correlations between two blocks of spins.

Next, we relate correlations with entanglement. We note that after the measurement process and for an initially pure state we end up with a pure state of two qubits. For such a state we have proven the following result [10,11]:
(1.ii) The entanglement of a pure two qubit state $|\psi_{ij}\rangle$ as measured by the concurrence is equal to the maximal correlation function:

$$C(\psi_{ij}) = \max_{\vec{a},\vec{b}} |\ Q^{ij}_{AB}(\psi_{ij})|. \tag{22}$$

Combining *(1.i)* and *(1.ii)* we know that for a given pure multi qubit state $|\psi\rangle$ and directions $\vec{a}, \vec{b}$ there always exists a measurement $\mathcal{M}$ such that:

$$Q^{ij}_{AB}(\psi) \leq \sum_s p_s Q^{ij}_{AB}(\psi_s) \leq \sum_s p_s C(\psi_s). \tag{23}$$

The term on the very right is equal to the average entanglement as measured by the concurrence $L^{\mathcal{M},C}_{ij}$, which trivially is a lower bound to the LE as defined by L^C_{ij}. Since the directions $\vec{a}, \vec{b}$ can be chosen arbitrarily, relation (23) holds in particular for directions maximizing $Q^{ij}_{AB}(\psi)$. Hence we can establish the desired lower bound on LE [10]:

(1.iii) Given a pure state $|\psi\rangle$ of N qubits, then the LE as measured by the concurrence is larger or equal than the maximal correlation:

$$L^{C}_{i,j}(\psi) \geq \max_{\vec{a},\vec{b}} |Q^{ij}_{AB}(\psi)|. \tag{24}$$

Making use of the basic properties of LE, presented in the previous section, we can immediately derive analogous bounds for some other variants of LE; for example,

$$L^{E_F}_{i,j}(\psi) \geq f(\max_{\vec{a},\vec{b}} |Q^{ij}_{AB}(\psi)|), \tag{25}$$

$$L^{\mathrm{POVM,E_E}}_{i,j}(\psi) \geq f(\max_{\vec{a},\vec{b}} |Q^{ij}_{AB}(\psi)|), \tag{26}$$

with f being the convex function defined in (5). We will see below that relation (26) can be generalized to spin-1 systems.

4.2.2. Higher Dimensional Spin Systems

We now try to extend the previous findings beyond spin-1/2 systems. First, we look for a generalized version of statement *(1.i)*. Unfortunately the techniques used in the proof for qubits seem to fail already for qutrits. Nevertheless, a generalization is still possible by changing a little bit the perspective. For this we embed a spin-S in a higher dimensional Hilbert space, being composed of $n \geq \log_2(2S+1)$ *virtual* qubits. Let us denote the $(2S+1) \times 2^n$ matrix governing this transformation by P.

In the case $2S+1 = 2^n$ the embedding is trivial and the situation becomes equivalent to the qubit case. Thus the result *(1.i)* can immediately be generalized, because local measurements on the virtual qubit systems can be chosen such that *(1.i)* holds.

In the case $2S+1 < 2^n$ a similiar argument applies if we allow for POVM measurements on the spin-S system. To be more precise, let us consider a mixed state ρ of three spin-S particles. The spin on which the measurement is performed (let us denote it with the index 3) is embedded in a 2^n dimensional system. The embedded state is then given by the transformation: $\rho' = (\mathbb{1}_{12} \otimes P_3^{\dagger})\rho(\mathbb{1}_{12} \otimes P_3)$. In the Hilbert space of the n virtual qubits one always finds local projective measurements $\{M_{\alpha_1 \ldots \alpha_n}\} = \{|\alpha_1\rangle\langle\alpha_1| \otimes \ldots \otimes |\alpha_n\rangle\langle\alpha_n|\}$ such that (the generalized version of) *(1.i)* holds for the state ρ'. In terms of the original state ρ this measurement in the 2^n dimensional space corresponds to a POVM measurement $\{PM_{\alpha_1 \ldots \alpha_n}P^{\dagger}\}$ on a spin-S system, because: $\sum_{\alpha_1 \ldots \alpha_n} PM_{\alpha_1 \ldots \alpha_n}P^{\dagger} = P\mathbb{1}_{2^n \times 2^n}P^{\dagger} = \mathbb{1}_{(2S+1)\times(2S+1)}$.

Thus one can generalize the result *(1.i)* to arbitrary spin dimensions in the following way:

(2.i) Given an arbitrary multi-spin state with connected correlation function Q^{ij}_{AB} between spins i and j for arbitrary operators S_A, S_B, then there always exists a local POVM measurement on the other spins such that this correlation does not decrease, on average.

We note that the lower bound in *(2.i)* can already be reached by applying *local* measurements on the virtual qubit system. Performing joint measurements (e.g. Bell measurements on pairs of qubits as shown in [12]) can lead to a con-

siderable enhancement of the average correlations. If in addition $2^n - (2S+1)$ joint measurements can be chosen such that they are orthogonal to the projector P, the resulting measurement on the spin-S system corresponds to a projective von-Neumann measurement.

The most difficult part is to establish a connection between correlations and entanglement for pure two spin states in analogy of *(1.ii)*. In the case of qubits we made explicitly use of the fact that the group $SU(2)$ is the covering group of $SO(3)$. Moreover the concurrence served as an entanglement measure, which was easy to handle. For higher spin dimensions we refer to the entropy of entanglement E_E (3) as a suitable entanglement measure for pure bipartite states. In the special case of qutrits we were able to show the following relation [11]:
(2.ii) The entanglement of a pure two qutrit state $|\psi_{ij}\rangle$ as measured by the entropy of entanglement can be lower bounded by:

$$E_E(\psi_{ij}) \geq f(\max_{A,B} |\ Q^{ij}_{AB}(\psi_{ij})|), \tag{27}$$

where f is the convex function (5) and S_A, S_B in Q^{ij}_{AB} (21) are operators, whose eigenvalues lie in the interval $[-1;1]$. Combining again *(2.i)* and *(2.ii)* we can formulate a bound on LE for spin-1 systems:
(2.iii) Given a pure state $|\psi\rangle$ of N qutrits, then the LE as measured by the entropy of entanglement and which allows for POVM's, is lower bounded by the maximum connected correlation function in the following way:

$$L^{\mathrm{POVM,E_E}}_{ij}(\psi) \geq f(\max_{A,B} |Q^{ij}_{AB}(\psi)|). \tag{28}$$

In summary, we have shown for pure qubit and qutrit states that connected correlation functions provide a lower bound on LE. This bound allows for two intriguing limiting case: (i) Entanglement and correlations may exhibit similiar behavior. (ii) Spins may be maximally entangled although they are uncorrelated in the classical sense. In the forthcoming sections we will present examples for both scenarios.

5. Numerical Computation of LE

In this section we review briefly the scheme, proposed in [11], for the numerical calculation of LE in ground states of 1D spin systems. Later we will show that this method can also be generalized to thermal states. Let us consider a chain of N spins of dimension $d = 2S+1$. The ground state of the system can be determined exactly by diagonalization of the Hamiltonian. However, the state is characterized by an exponential amount of parameters as a function of N, thus limiting the exact treatment of the problem to relatively small system sizes.

As an alternative approach we can start out with an approximation of the exact ground state in terms of the the so-called matrix product states (MPS) [23,22]:

$$|\psi_{\mathrm{MP}}\rangle = \sum_{s_1,\ldots,s_N=1}^{d} \mathrm{Tr}(A_1^{s_1}\ldots A_N^{s_N})|s_1,\ldots,s_N\rangle. \tag{29}$$

Here the state is described by N matrices $A_i^{s_i}$ of dimension D. We note that the MPS (29) is written in the computational basis and accounts for periodic boundary conditions (PBC). It has been shown [23,38] that MPS appear naturally in the context of the density-matrix renormalization group (DMRG) method [39]. In [11] we propose a scheme to obtain a translationally invariant MPS representation of the ground state from the DMRG algorithm for an infinite chain. Alternatively, one can optimize the wavefunction directly in the MPS representation. In contrast to standard DMRG, this approach makes it possible to calculate ground state properties of systems with PBC with high accuracy [40]. Moreover, it can be used to simulate the time evolution of pure and mixed states, and thus allows for the calculation of thermal states [51].

Once we have calculated the ground state of a system in MPS form (29), it is easy to obtain the outcomes of local measurements, as required in the definition of LE. However, the calculation of the average entanglement (14) involves the summation of an exponential number of terms. This problem can be circumvented by using the Monte Carlo (MC) method [24], which selects the most probable states from a given probability distribution. Hence, we have an efficient method for calculating the average entanglement (14) for a given measurement basis. We are still facing the problem of finding the measurement strategy, which maximizes the average entanglement. A good guess for the optimal basis can typically be found using exact diagonalization for small system sizes, followed by numerical maximization of the average entanglement. Alternatively, the optimal basis can be extracted directly from the MPS matrices A_i in (29) [12,11]. For the models we studied our numerical analysis indicates that the optimal basis appears to be independent of the system size. Hence exact diagonalization and numerical optimization for small systems usually provides the most efficient way to find the optimal measurement strategy also for large systems. Moreover, it turns out that, typically, the optimal strategy is to measure all sites in the same basis.

6. Examples of LE in Spin-1/2 Models

Let us now apply the concept of LE to quantify the localizable ground state entanglement of various spin-1/2 models. After some general considerations we compute the LE as measured by the concurrence $L^C_{i,j}$ numerically for two specific examples.

6.1. General Considerations

We consider spin-1/2 Hamiltonians of the form

$$H = -\sum_{i,j}\sum_{\alpha=x,y,z} \gamma_\alpha^{ij}\sigma_\alpha^i\sigma_\alpha^j - \sum_i \gamma^i\sigma_z, \tag{30}$$

with parity symmetry, $[H, \Pi_z] = 0$ and $\Pi_z := \otimes_{i=1}^{N} \sigma_z^i$. Extensive numerical calculations on systems of up to 20 qubits showed that our lower bound is always close to the LE as measured by concurrence $L_{i,j}^C$, and typically is exactly equal to it: this is surprising and highlights the thightness of the given lower bound. Note also that whenever parity symmetry is present, the upper and lower bound are given as follows [10]:

$$\max\left(|Q_{xx}^{ij}|, |Q_{yy}^{ij}|, |Q_{zz}^{ij}|\right) \leq L_{i,j}^C \leq \frac{\sqrt{s_+^{ij}} + \sqrt{s_-^{ij}}}{2}, \tag{31}$$

$$s_\pm^{ij} = \left(1 \pm \langle \sigma_z^i \sigma_z^j \rangle\right)^2 - \left(\langle \sigma_z^i \rangle \pm \langle \sigma_z^j \rangle\right)^2.$$

The fact, that the lower bound is usually tight, can also be derived from the numerical observation that for Hamiltonians of the form (30) measurements in the (standard) σ_z-basis ($\mathcal{M} = \mathcal{Z}$) yield in most cases the optimal result. Expanding the ground state in that basis, $|\psi\rangle_0 = \sum c_{i_1 \ldots i_N} |i_1 \ldots i_N\rangle$, it is straightforward to show that e.g. in the case of constant (site independent) coupling γ_x and γ_y the ground state energy is minimized if all expansion coefficients $c_{i_1 \ldots i_N}$ have the same sign. This guarantees, together with the parity symmetry, that the average entanglement $L_{i,j}^{\mathcal{Z},C}$ for measurements in the standard basis is equal to either the $x-x$ or $y-y$ correlation. To be more precise we distinguish the following cases:

$$(\gamma_x^{ij} - \gamma_y^{ij})(\gamma_x^{ij} + \gamma_y^{ij}) \geq 0: \quad L_{i,j}^{\mathcal{Z},C} = |\langle \sigma_x^i \sigma_x^j \rangle|, \tag{32a}$$

$$(\gamma_x^{ij} - \gamma_y^{ij})(\gamma_x^{ij} + \gamma_y^{ij}) \leq 0: \quad L_{i,j}^{\mathcal{Z},C} = |\langle \sigma_y^i \sigma_y^j \rangle|, \tag{32b}$$

where the conditions refer to all sites i and j of the chain. Most of the prominent spin Hamiltonians studied in literature, like the Heisenberg, XY or XXZ model etc., trivially fulfill one of the conditions (32), because their coupling coefficients are site-independent. Hence, measurements in the standard basis would yield localizable quantum correlations that are completely determined by classical correlations.

6.2. Ising Model

As an illustration, let us now discuss the LE of the Ising model in a transverse magnetic field ($\gamma_\alpha^{ij} = \lambda \delta_{\alpha,x} \delta_{j,i+1}; \gamma^i = 1$ in (30)), which has been solved exactly [41] and exhibits a quantum phase transition at $\lambda = 1$. In this case, the maximal connected correlation function is always given by Q_{xx}, which thus yields the best lower bound on LE. Numerical optimization for a finite chain indicates that the standard basis is indeed the optimal one and thus the lower bound is equal to $L_{i,j}^C$ [10]. We checked analytically, using perturbation theory, that for an infinite chain this numerical result is indeed true. However, for a spin distance $n = |i - j|$ one has to go to n-th order perturbation theory, limiting this analytical treatment to rather small n.

Consequently, we can use exact results for the connected correlation function Q_{xx} [41] to completely characterize the behavior of the LE in the Ising chain. The Ising system is therefore also an ideal candidate for testing the performance of

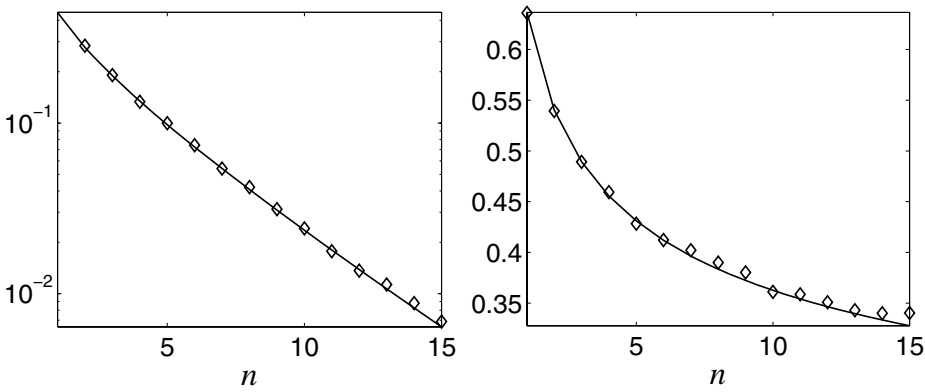


Figure 2. Calculation of the $L^C_{i,i+n}$(diamonds) for the ground state of the Ising chain in a transverse field ($\gamma^{ij}_\alpha = \lambda\delta_{\alpha,x}\delta_{j,i+1}; \gamma^i = 1$ in (30)) as a function of the spin distance n. For comparison we plot the exact result [41] for the correlation function $Q^{i,i+n}_{xx}$. Left: $\lambda = 0.8$, exponential decrease; Right: critical point $\lambda = 1$, power law decrease ($\sim n^{-1/4}$); Numerical parameters (see Sect. 5): $N = 80$, $D = 16$, MC sweeps $M = 20,000$.

our numerical method outlined in Sect. 5. In Fig. 2 we plot $L^C_{i,i+n}$ and $Q^{i,i+n}_{xx}$ as a function of the spin distance n for a chain with $N = 80$ sites.

For $\lambda < 1$, the LE decreases exponentially with n, and the entanglement length is finite. At the quantum critical point $\lambda = 1$, the behavior of the LE changes drastically, because it suddenly decreases as a power law, $L^C_{i,i+n} \sim n^{-1/4}$, thus leading to a diverging entanglement length ξ_E. In Fig. 2 we observe that the MC method becomes less accurate at the critical point. As we will see later, one reason is that the statistical error due to entanglement fluctuations becomes rather large at the critical point (see Fig. 3). Another (systematic) error might be induced by the single-spin-flip dynamics used to create the trial state. Better results for the critical region could possibly be achieved by applying the Wolff algorithm [43]. Here, a cluster of spins depending on their spin orientation is flipped, which accounts for the formation of domains.

In [10] it was shown that for the case $\lambda > 1$ we also get $\xi_E = \infty$, since the LE saturates to a finite value given by $M_x^2 = 1/4(1 - \lambda^{-2})^{1/4}$. Indeed, the ground state is then close to the GHZ-state. In a more realistic setup, however, the parity symmetry of the Ising Hamiltonian will be broken by a perturbation and the ground state for large coupling will also be separable, as it is given by a superposition of two GHZ states with different parity [42].

Let us now study the behavior of entanglement fluctuations $\delta L^C_{i,j}$ inherent to the statistical definition of LE (see Def. (19)) as a function of the coupling λ. In Fig. 3 we plot $\delta L^C_{i,j}$ for different parameters n and N.

The maximum of the fluctuations is always located in the vicinity of the critical point $\lambda = 1$ and gets shifted to larger λ values with increasing N. Thus the increasing entanglement fluctuations reflect very well the increasing complexity of the wavefunction close to the critical region. The location of the maximum λ_m in the thermodynamic limit ($N \to \infty$) apparently depends on the distance n of the two spins. For nearest neighbors ($n = 1$) we observe that the maximum of

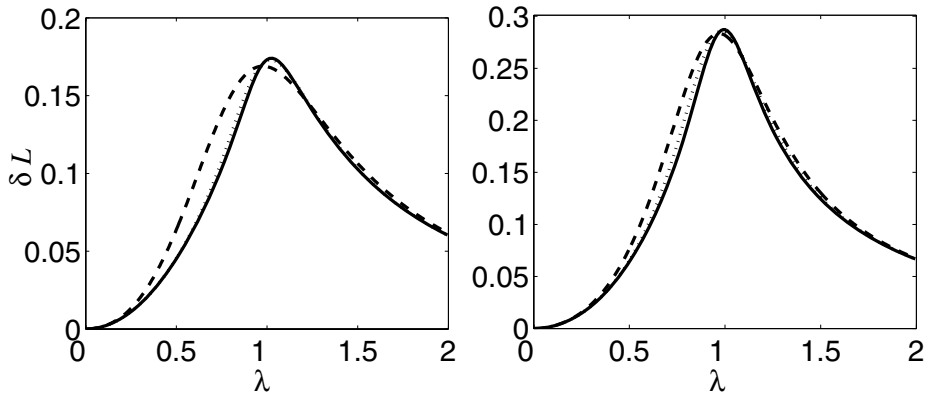


Figure 3. Exact calculation of entanglement fluctuations $\delta L^C_{i,i+n}$ (19) as a function of the coupling parameter λ for a finite Ising chain in a transverse magnetic field with PBC. Left: distance $n = 1$, $N = 6$ (dashed), $N = 12$ (dotted), $N = 16$ (solid); Right: distance $n = 4$, $N = 9$ (dashed), $N = 12$ (dotted), $N = 16$ (solid).

$\delta L^C_{i,i+1}$ is somewhat shifted to the right of the critical point ($\lambda_m \approx 1.025$). For all distances $n > 1$, however, our numerical calculations show that the maximum is positioned at $\lambda < 1$ but becomes asymptotically close to the critical point with increasing n (for $n = 4$ see Fig. 3). Furthermore, in Fig. 3 we see that the absolute value of the maximum increases with n and becomes comparable with L^C_{ij} itself [44]. The strong fluctuations inherent to the Ising model lead to large statistical errors in the numerical calculation of LE using Monte Carlo. The errors become even more pronounced for the calculation of the fluctuations. This is the main reason, why we have restricted ourselves here to exact calculations for a small system with PBC. However, we confirmed that the data for $N = 16$ represents the behavior in the large N limit reasonably well and no qualitative changes occur.

6.3. XXZ Model

Let us now turn to the discussion of another exactly solvable 1D spin system, the so called XXZ model [45]. This model not only appears in condensed matter physics in the context of ferro- or antiferromagnetic materials. Recently it has been shown that it can also effectively describe the physics of ultra cold atoms in a deep optical lattice [50]. The Hamiltonian can be written as

$$H_{\mathrm{XXZ}} = -\sum_i \left[\sigma_x^i\ \sigma_x^{i+1} + \sigma_y^i\ \sigma_y^{i+1} + \Delta\ \sigma_z^i\ \sigma_z^{i+1} + \frac{h}{J}\sigma_z^i\ \right], \tag{33}$$

where we have introduced two dimensionless parameters, which can be varied independently: the anisotropy Δ and the magnetic field h/J in units of the exchange coupling.

We consider the case of zero magnetic field and study the LE of the ground state as a function of Δ in the vicinity of the antiferromagnetic XXX point. At the critical point $\Delta = -1$ the ground state undergoes a Kosterlitz-Thouless [20] quantum phase transition. In Fig. 4 we have calculated numerically the localizable

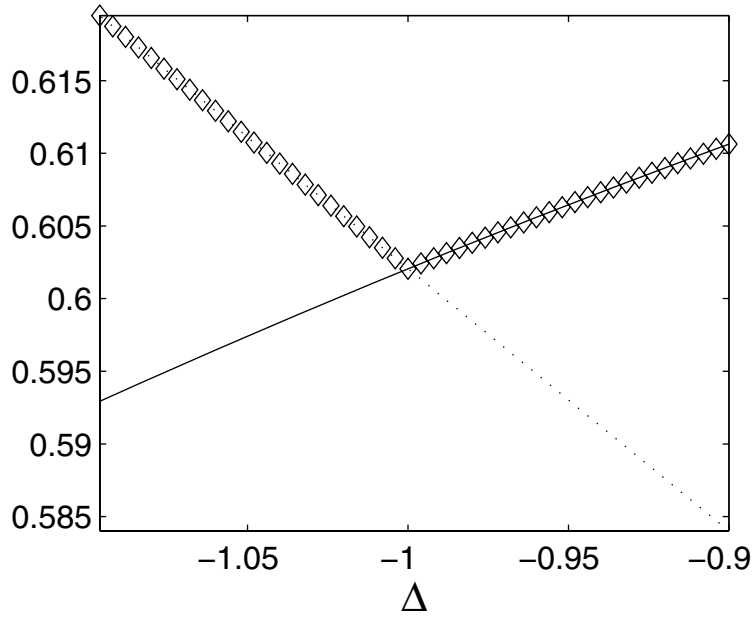


Figure 4. Calculation of $L^C_{i,i+1}$ (diamonds) and the lower bounds Q^{ii+1}_{xx} (solid) and Q^{ii+1}_{zz} (dotted) as a function of the anisotropy Δ for the XXZ model (33) with zero field h/J. At the critical point $\Delta = -1$ the LE exhibits a cusp, in contrast to the monotonic behavior of the correlation functions. The calculation is performed using exact diagonalization and numerical optimization for a chain of length $N = 10$ with periodic boundary conditions.

entanglement $L^C_{i,i+1}$ for nearest neighbors and the corresponding lower bounds $Q^{i,i+1}_{xx}$ and $Q^{i,i+1}_{zz}$.

One sees that $L^C_{i,i+1}$ is equal to the maximum correlation function. However, at the critical point the maximum correlation function changes, due to the crossing of $Q^{i,i+1}_{xx}$ and $Q^{i,i+1}_{zz}$, thus leading to a cusp in the LE. Hence the quantum phase transition is characterized by a discontinuity in the first derivative of $L^C_{i,i+1}$. This result is remarkable, because for this Kosterlitz-Thouless transtion the ground state energy (i.e nearest neighbor correlation functions) and all of its derivatives are continous [47]. As shown in [48,49], the concurrence (or the derivative) of the reduced density matrix $\rho_{i,i+1}$ also does not exhibit a discontinuity at the critical point. This result is expected, because the concurrence is a function of one-particle and two-particle correlation functions. We further note that, according to our numerical analysis the reason for the cusp in $L^C_{i,i+1}$ is that the optimal measurement basis changes at the critical point abruptly from the σ_z- to the σ_x-basis.

So far this is a purely numerical finding for a finite dimensional system. However, given that the average entanglement can be maximized by applying the same unitary transformation on all spins, one can rigorously show that indeed a cusp in the LE must occur exactly at the critical point and independently of the size N. The argument goes at follows: At the antiferromagnetic XXX point the Hamiltonian (33) possesses $SU(2)$ symmetry. This means that any measurement basis

yields the same LE. In particular, we know from (32a) that for measurements in the standard basis the LE is equal to the correlation function $Q_{xx}^{i,i+1}$. At the critical point we thus have: $L_{i,i+1}^{C} = Q_{zz}^{i,i+1} = Q_{xx}^{i,i+1}$. Since connected correlation functions yield a lower bound to $L_{i,i+1}^{C}$ (see Sect. 4) the localizable entanglement $L_{i,i+1}^{C}$ must exhibit a cusp at the critical point, where $Q_{zz}^{i,i+1}$ and $Q_{xx}^{i,i+1}$ cross.

In summary our discussion of the LE in various spin-1/2 models has shown in which parameter regimes these systems can be used for e.g. localizing long-range entanglement as indicated by the entanglement length. We have further seen that the study of LE as well as the entanglement fluctuation provides a valuable tool for detecting and characterizing quantum phase transitions. In addition our numerical results indicate that the ground state entanglement of spin-1/2 Hamiltonians with two-spin nearest neighbor interactions is typically very well described by the maximum correlation function. We will now see that this observation is not necessarily true for spin-1 systems.

7. LE in the Spin-1 Antiferromagnetic Heisenberg Chain

We study the ground state entanglement of the (generalized) antiferromagnetic Heisenberg chain:

$$H_{AF} = \sum_{i=1}^{N-1} \left[\vec{S}_i \cdot \vec{S}_{i+1} - \beta (\vec{S}_i \cdot \vec{S}_{i+1})^2 \right], \tag{34}$$

which includes a biquadratic term. In a recent work [52] it has been demonstrated that quantum Hamiltonians of this kind can be implemented with ultra cold atoms trapped in an optical lattice potential.

7.1. General Considerations

For the Heisenberg antiferromagnet (AF) ($\beta = 0$) it has been shown by Haldane [53] that in the case of half-integer spins the spectrum is gapless in the thermodynamic limit, and thus the correlation length of the ground state is infinite. For integer spins, however, an energy gap emerges, resulting in a finite correlation length. Let us now investigate whether a similiar connection holds for the LE and the corresponding entanglement length.

¿From the lower bound (24) it follows that the predicted infinite correlation length in spin-1/2 systems (with integer $\log_2(2S+1)$) automatically implies a diverging entanglement length. However, in the case of integer spins the correlation length is finite. Hence the lower bound (25) for spin-1 systems includes the intriguing possibility that correlations and entanglement may exhibit a completely different behavior. Indeed, for gapped spin-1 systems an example for this has already been found [12]. At the AKLT [21] point ($\beta = -1/3$) the entanglement length of the ground state diverges, although the correlations decrease exponentially. Since the AKLT model is closely related to the Heisenberg AF, one might expect that both systems show qualitatively the same behavior.

7.2. Numerical Study

Let us now have a closer look on the spin-1 Heisenberg AF. Since in this case we cannot resort to an analytical solution as for the AKLT, we have to rely on numerical methods.
We start our analysis by performing exact diagonalizations for an open chain of up to 10 sites. At the endpoints we couple to $S = 1/2$ spins thus making sure that the system is in the singlet ground state. We are interested in the LE between the endpoints of the chain. Since the end spins are represented by qubits we can still refer to the LE as measured by the concurrence: $L^C_{1,N}$. Our numerical analysis shows that the optimal measurement basis is given by the same local unitary transformation:

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{bmatrix} , \tag{35}$$

as for the AKLT. This strategy produces a maximally entangled state between the end spins ($L^C_{1,N} = 1$). This surprising result can be understood from the analytical study of the AKLT model in [12]. In a singlet valence bond picture measurements in the basis (35) can be interpreted as Bell measurements on a (virtual) spin-1/2 system, which lead to entanglement swapping.

Let us now investigate, whether this effect depends on the choice of the boundary condition or the number of sites. For this purpose we apply our numerical method outlined in Sect. 5. The LE in its variant $L^{E_E}_{ij}$ (see (3)) is computed for a chain with PBC and a large number of sites ($N = 80$).

In Fig. 5 we see that the LE saturates at a finite value $L^{E_E}_{i,i+n} \to 0.960 \pm 0.003$ for large n, whereas the correlations decrease exponentially. This demonstrates that the ground state of the antiferromagnetic spin-1 Heisenberg chain could be used to distribute EPR-like entanglement over arbitrary distances by performing local operations on the intermediate spins. As mentioned in Sect. 3 this result might be particularly interesting in the context of quantum repeaters.
Let us now come back to Haldane's result for the Heisenberg AF stated in the beginning of this section. Our numerical study of the spin-1 case might give a first indication that, unlike the correlation length, the entanglement length is infinite for both half-integer and integer spins.

7.3. Hidden Order and String Order Parameter

Our numerical results show that the ground state of the spin-1 Heisenberg AF exhibits long range order in terms of the localizable entanglement. We have also seen that this long range order is not reflected in the behavior of the two-particle correlation functions. However, one can define a multiparticle correlation function, the so called *string order correlation function* [54], which detects this *hidden order* in the ground state. The string order correlation function has been argued to be of topological nature and is defined as

$$Q^{i,i+n}_{so} = \langle S^i_z \left[\otimes^{i+n-1}_{k=i+1} R_k \right] \otimes S^{i+n}_z \rangle, \tag{36}$$

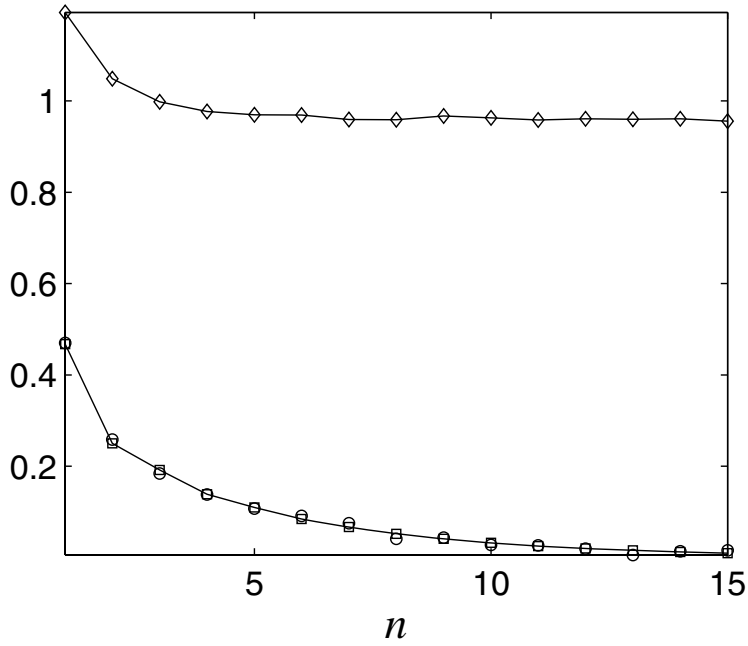


Figure 5. Calculation of $L^{E_E}_{i,i+n}$ (diamonds) for the ground state of the antiferromagnetic spin-1 Heisenberg chain with $N = 80$ sites as a function of the spin distance n. For comparison we plot the correlation function $Q^{i,i+n}_{xx}$ computed directly from the MPS (squares) and using Monte Carlo (circles). Numerical parameters (see Sect. 5): $D = 16$, MC sweeps $M = 20,000$.

with $R_k = \exp\left(i\pi S^k_z\right)$. A non-vanishing *string order parameter*, $\xi_{so} := \lim_{n\to\infty} Q^{i,i+n}_{so}$, indicates the presence of long range (hidden) order. As an obvious generalization of the string order correlation function to arbitrary models, let us define a connected version in a variational way. Consider the set of all observables $\{\hat{O}\}$ with bounded spectrum $-\mathbb{1} \leq \hat{O} \leq \mathbb{1}$. We define the *connected string order correlation function* $Q^{i,i+n}_{cso}$ (and the related parameter ξ_{cso}) for a given translational invariant state as

$$Q^{i,i+n}_{cso} = \max_{-\mathbb{1}\leq\hat{O}_1,\hat{O}_2\leq\mathbb{1}} \langle \hat{O}^i_1 \left[\otimes^{i+n-1}_{k=i+1} \hat{O}^k_2\right] \hat{O}^{i+n}_1 \rangle_c. \tag{37}$$

Here $\langle A_1 A_2 \ldots A_n \rangle_c$ denotes the connected n-point correlation function, which can be defined in a recursive way:

$$\langle A_1\rangle = \langle A_1\rangle_c, \tag{38}$$

$$\langle A_1A_2\rangle = \langle A_1\rangle_c\langle A_2\rangle_c + \langle A_1A_2\rangle_c, \tag{39}$$

$$\begin{aligned}\langle A_1A_2A_3\rangle = {} & \langle A_1\rangle_c\langle A_2\rangle_c\langle A_3\rangle_c + \langle A_1\rangle_c\langle A_2A_3\rangle_c \\ & + \langle A_2\rangle_c\langle A_1A_3\rangle_c + \langle A_3\rangle_c\langle A_1A_2\rangle_c \\ & + \langle A_1A_2A_3\rangle_c,\end{aligned} \tag{40}$$

$$\vdots$$

Note that the connected part assures that $Q_{cso}^{i,i+n}$ measures a nonlocal correlation, and that the string order parameter of the AKLT-ground state is indeed recovered by this definition.

It has been verified numerically that ξ_{so} is finite for the ground state of the spin-1 Heisenberg AF. This fact can rigorously be proven for the related AKLT-ground state [21]. For this state it was further shown that the LE saturates as well with the spin distance n [12]. Hence one might expect a connection between the existence of long range order in the entanglement and long range order indicated by the string order parameter. However, one can find examples for which this connection does not hold.

For instance, in [12] it has been shown that already an infinitesimal deformation of the AKLT model leads to an exponentially decreasing LE, whereas ξ_{cso} stays finite. On the other hand, ground states exist that exhibit a diverging entanglement length but vanishing ξ_{cso}. A simple example can be found in the class of MPS (29) defined on qubits ($d = 2$) and with qubit bonds ($D = 2$). Note that all these MPS are guaranteed to be ground states of some local Hamiltonians. Furthermore, for MPS with qubit bonds the string order parameter and the LE can easily be computed analytically (see [12] and [11]). In particular, let us study the state defined by

$$A^1 = \sigma_z + \sigma_y \qquad A^2 = \sigma_z - i\mathbb{1}. \tag{41}$$

The entanglement length can easily be proven to be infinite. A necessary condition for ξ_{cso} to be nonzero is that there exists a unitary operator $\hat{O}_2$ for which the largest eigenvalue of $E_{\hat{O}_2}$ has the same magnitude as the maximal eigenvalue of $E_{\mathbb{1}}$. For the example given, this is impossible, hence providing an example of a ground state with a diverging entanglement length but no long range hidden order.

8. Mixed States

In this section we apply the concept of LE to characterize the entanglement of multipartite mixed states. Note that the definition of LE (15) already includes the possibility of having a mixed state ρ. This implies that the states $\rho_{\{s\}} = \langle\{s\}|\rho|\{s\}\rangle$ after the measurements are also mixed. Hence we refer in the following to the LE as measured by the negativity: L_{ij}^N. In order to provide a tool for the computation of L_{ij}^N let us now generalize the numerical method, outlined in Sect. 5, to mixed states.

8.1. Numerical Method

The key point is to find a representation of a mixed state in terms of low dimensional matrices A_k, analogous to the MPS (29). This problem has been considered recently in [51]. There the concept of MPS is generalized to matrix product density operators (MPDO), which are defined as

$$\rho = \sum_{s_1,s_1',\ldots,s_N,s_N'=1}^{d} \mathrm{Tr}(M_1^{s_1,s_1'} \ldots M_N^{s_N,s_N'}) \times |s_1,\ldots,s_N\rangle\langle s_1',\ldots,s_N'|, \tag{42}$$

where $M_k^{s_k,s_k'}$ are $D_k^2 \times D_{k+1}^2$ matrices. They can be decomposed as

$$M_k^{s,s'} = \sum_{a=1}^{d_k} A_k^{s,a} \otimes (A_k^{s',a})^*. \tag{43}$$

The state ρ can be purified into a MPS by including ancilla states $\{|a_k\rangle\}$ of dimension d_k:

$$|\Psi\rangle = \sum_{s_1,\ldots,s_N} \sum_{a_1,\ldots,a_N} \mathrm{Tr}\left(\prod_{k=1}^{N} A_k^{s_k,a_k}\right) |s_1 a_1,\ldots,s_N a_N\rangle. \tag{44}$$

In [51] a method has been introduced that allows one to determine the matrices A_k iteratively for a given Hamiltonian and temperature.
Starting from the mixed state in MPDO representation (42) the LE can be computed along the same lines, using the MC method, as for pure states. However, we note that the matrices M_k in (42) have dimension $D^2 \times D^2$, compared to the $D \times D$ matrices A_k for pure states. Thus in the case of PBC the computation time for the MC part scales at least with $d\ D^5$ for mixed states, but only with D^3 for pure states.
Let us also comment briefly on the optimal measurement basis. In contrast to pure states, the optimal basis for mixed states can no longer be deduced directly from the matrices M_k. Alternatively a good guess for the best measurement strategy can be found by exact diagonalization of the Hamiltonian for small N, followed by numerical optimization of LE.

8.2. Example: AKLT Model

Next we apply this numerical method to a specific example. Interesting candidates can be found in gapped spin-1 systems, like the AKLT model or the Heisenberg AF (see (34)). We have seen that the ground state of these two models exhibits an infinite entanglement length [12]. Thus the natural question arises, to what extent this feature holds for small but finite temperatures and how the entanglement length scales with temperature.
For our numerical study we choose the AKLT model. The reason is simply that,

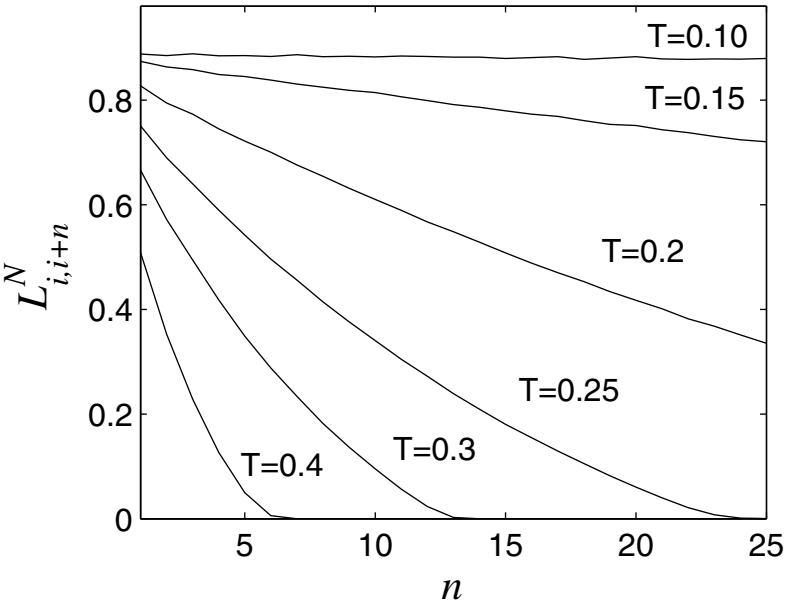


Figure 6. Calculation of the LE as given by $L^N_{i,i+n}$ for the AKLT model as a function of the spin distance n and for various temperatures T. We have chosen a chain with OBC and $N = 50$ sites. We note that almost identical data can be obtained for $N = 20$, indicating that our results are already close to the thermodynamic limit. Numerical parameters (see also Sect. 5): matrix dimension $D = 10$, MC sweeps $M = 5,000$.

here, mixed states can be approximated rather well by matrices with dimensions as small as $D \approx 10$, even for very small temperatures. This makes the computation of LE much more efficient compared to the Heisenberg AF. Efficiency is also the reason for choosing OBC.

Let us now discuss the optimal measurement strategy. First of all we point out, that with OBC the ground state of the AKLT is four-fold degenerate [55]. Thus for $T \to 0$ the density matrix is an equal mixture of these four states, which strongly reduces the LE compared to e.g. the singlet ground state studied in [12]. On the other hand it is known that the degeneracy results only from the end spins of the chain. Thus, one can strongly reduce this boundary effect by choosing the two spins, i and $i + n$, to be far away from the boundaries. For this situation we found that the optimal measurement scheme for LE is given by measurements in the U-basis (35) on the spins between sites i and $i + n$, and in the standard basis everywhere else. This result is not very surprising in terms of the valence bond picture in [12]. Entanglement swapping is only needed between the two spins of interest, whereas the effect of the degeneracy can be minimized by measuring the outer spins in the standard basis. Using this strategy we plot in Fig. 6 the LE, as given by $L^N_{i,i+n}$, depending on the spin distance n for various temperatures T. The temperatures are chosen to be of the order of the enery gap [56].

The data indicates an exponential increase of the entanglement length $\xi_E \sim e^{\alpha/T}$ with $\alpha \approx 0.8$, thus leading smoothly to an infinite entanglement length

at zero temperature. This behavior is not unexpected for a 1D system from the perspective of the Mermin-Wagner theorem [57]. However, it is not clear whether this theorem is really applicable to phase transitions in terms of LE. It is more inspiring to treat this problem on the basis of *projected entangled-pair states* [58]. In this picture a finite temperature phase transition for LE could possibly occur for two or more dimensions. We note that recently such a transition has been shown to exist for 3D cluster states [15].
Finally we would like to point out that, although the entanglement length of the AKLT model is finite for $T > 0$, it can still be considerably large for sufficiently low temperatures $T \lesssim 0.2$. Thus for practical purposes this system might still be useful, e.g. for quantum repeater setups.

9. Conclusion

In this article we have reviewed the concept and basic properties of LE, as defined in [10] and [11]. In particular, we have stated bounds on the LE in terms of classical connected correlation functions. Moreover we have briefly presented a numerical method, applicable to both pure and mixed states, that allows to calculate the LE efficiently even for large 1D spin systems. In future work we are planning to generalize this numerical scheme to two and higher dimensional spin systems based on the method [58]. We have applied the concept of LE to study the entanglement properties of ground states of various 1D spin models. The results can be summarized as follows:

The LE, as well as the entanglement fluctuations, exhibit characteristic features at a quantum phase transition. This result is a direct consequence of the numerical observation, that, for ground states of spin-1/2 systems with two-spin interactions, the lower bound is typically tight, i.e. the LE is completely characterized by the maximal connected correlation function. However, connected correlation functions give only a coarse grained picture for the LE. As an example, we have shown for the spin-1 Heisenberg AF that the entanglement length diverges, whereas the correlation length is finite. This numerical result might suggest that, in terms of LE and opposed to connected correlation functions, no distinction has to be made between the scaling behavior of integer and half-integer spin Heisenberg AF's. To confirm this idea, further studies of the integer spin case are desired.
We further note that preliminary results indicate that the entanglement features of the Heisenberg AF might hold qualitatively for the whole class of gapped spin-1 models defined by (34). This finding would allow one to distinguish the Haldane phase from the dimerized phase based on the scaling of LE. Let us also mention in this context that gapped spin-1 systems also appear to be ideal candidates to look for quantum phase transitions, being detected solely by the entanglement length and not the correlation length (like for the generalized AKLT model in [12]).

In order to illustrate that our numerical method works also for mixed states we computed the LE of the AKLT model for finite temperatures. We found that the

entanglement increases exponentially with the inverse temperature. This smooth behavior indicates the absence of a phase transition in terms of LE for the 1D case. However, we have reason to believe that such a phase transiton might occur for the 2D system.

10. Acknowledgements

We acknowledge helpful discussions with J.-J. García-Ripoll, who also provided the numerical program for computing MPDO's in the case of finite temperature. We also thank D. Loss for suggesting to address the issue of entanglement fluctuations. This work was supported in part by EU IST projects (RESQ and QUPRODIS), the DFG (SFB 631), the "Kompetenznetzwerk Quanteninformationsverarbeitung" der Bayerischen Staatsregierung, the DGS under contract BFM 2003-05316-C02-01 and the Gordon and Betty Moore Foundation (the Information Science and Technology Initiative, Caltech).

References

[1] C.H. Bennett et al., Phys. Rev. Lett. **70**, 1895 (1993).
[2] J. Preskill, J. Mod. Opt. **47**, 127 (2000).
[3] P. Zanardi and X. Wang (2002), J. Phys. A **35**, 7947 (2002).
[4] A. Osterloh, L. Amico, G. Falci, and R. Fazio, Nature **416**, 608 (2002).
[5] T. J. Osborne, and M. A. Nielsen, Phys. Rev. A **66**, 032110 (2002).
[6] G. Vidal, J.I. Latorre, E. Rico, and A. Kitaev, Phys.Rev.Lett. **90**, 227902 (2003); P. Calabrese, J. Cardy, J. Stat. Phys. P06002 (2004).
[7] D. Gunlycke, V. M. Kendon, V. Vedral, and S. Bose, Phys. Rev. A **64**, 042302 (2001).
[8] G. Ortiz, R. Somma, H. Barnum, E. Knill, L. Viola, quant-ph/0403043; R. Somma, G. Ortiz, H. Barnum, E. Knill, L. Viola, quant-ph/0403035;
[9] T. Roscilde, P. Verrucchi, A. Fubini, S. Haas, and V. Tognetti, Phys. Rev. Lett. **93**, 167203 (2004).
[10] F. Verstraete, M. Popp, and J.I. Cirac, Phys. Rev. Lett. **92**, 027901 (2004).
[11] M.Popp, F. Verstraete, M.A. Martin-Delgado, and J.I. Cirac, Phys. Rev. A **71**, 042306 (2005).
[12] F. Verstraete, M.A. Martin-Delgado, and J.I. Cirac, Phys. Rev. Lett. **92**, 087201 (2004).
[13] B.-Q. Jin and V.E. Korepin, Phys. Rev. A **69**, 062314 (2004);
[14] J.K. Pachos, M.B. Plenio, Phys. Rev. Lett. **93**, 056402 (2004); O.F. Syljuasen, Phys. Lett. A **322**, 25 (2004);W. Dür, L. Hartmann, M. Hein, H. J. Briegel, quant-ph/0407075; H. Fan, V. Korepin, V. Roychowdhury, quant-ph/0406067; V. Subrahmanyam, A. Lakshminarayan, quant-ph/0409048.
[15] Robert Raussendorf, Sergey Bravyi, Jim Harrington, quant-ph/0407255.
[16] M.B. Hastings, Phys. Rev. Lett. **93**, 140402 (2004).
[17] H.J. Briegel et al, Phys. Rev. Lett. **81**, 5932 (1998); W. Dür, H.-J. Briegel, J.I. Cirac, and P. Zoller, Phys. Rev. A **59**, 169 (1999).
[18] D.M. Greenberger, M. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, M. Kafatos, Ed.(Kluwer, Dordrecht 1989) pp. 69-72.

[19] H.J. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001).
[20] see e.g. S. Sachdev, *Quantum Phase Transitions*, (Cambridge University Press, Cambridge, 1999).
[21] I. Affleck, T. Kennedy, E.H. Lieb, and H. Tasaki, Commun. Math. Phys. **115**, 477 (1988).
[22] M. Fannes, B. Nachtergaele and R.F. Werner, Comm. Math. Phys. **144**, 443 (1992).
[23] S. Ostlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995).
[24] N. Metropolis, A.W. Rosensbluth, M.N. Rosensbluth, A.H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[25] For pure four qubit states it has been shown in [37], that in the case of projective measurements the LE cannot be increased by classical communication. In addition, numerical studies for the same system indicated that POVM measurements lead only to a very small increase of LE, compared to projective measurements.
[26] C.H. Bennett, H.J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A. **53**, 2046 (1996).
[27] W.K. Wootters, Phys. Rev. Lett. **80** 2245 (1998).
[28] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin,and W. K. Wooters, Phys. Rev. A. **54**, 3824 (1996); C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin, and W. K. Wooters, Phys. Rev. Lett. **76**, 722 (1996);
[29] G. Vidal, R.F. Werner, Phys. Rev. A **65**, 032314 (2002).
[30] M. Horodecki, P. Horodecki, R. Horodecki, Phys. Lett. A **223**, 1 (1996).
[31] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).
[32] M. Horodecki, P. Horodecki, R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
[33] P. Horodecki, M. Horodecki, R. Horodecki, Phys. Rev. Lett. **82**, 1056 (1999).
[34] D. Aharonov, Phys. Rev. A 62, 062311 (2000).
[35] J.I. Cirac, P. Zoller, H.J. Kimble, and H. Mabuchi, Phys. Rev. Lett. **78**, 3221 (1997); S.J. van Enk, J.I. Cirac, P. Zoller, Science **279**, 205 (1998); B. Kraus and J.I. Cirac, Phys. Rev. Lett. 92, 013602 (2004).
[36] D. P. DiVincenzo, C. A. Fuchs, H. Mabuchi, J. A. Smolin, A. Thapliyal, and A. Uhlmann, quant-ph/9803033
[37] T. Laustsen, F. Verstraete, and S. J. van Enk, Quantum Information and Computation **3**, 64 (2003).
[38] J. Dukelsky, M.A. Martin-Delgado, T. Nishino, G. Sierra, Europhys.Lett. **43**, 457 (1997).
[39] S.R. White, Phys. Rev. Lett. **69**, 2863 (1992); S.R. White, Phys. Rev. B **48**, 10345 (1993).
[40] F. Verstraete, D. Porras and J.I. Cirac, Phys. Rev. Lett **93**, 227205 (2004).
[41] P. Pfeuty, Ann. Phys. **57**, 79 (1970).
[42] V. Murg and J.I. Cirac, quant-ph/0309026.
[43] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).
[44] We note that for $\lambda < 1$ the ratio $\delta L^C_{i,j}/L^C_{i,j}$ can be much larger than one.
[45] see e.g., M. Takahashi, *Thermodynamics of One-Dimensional Solvable Models*, (Cambridge University Press, Cambridge, 1999).
[46] H.A. Bethe, Z. Physik **71**, 205 (1931).
[47] C.N. Yang and C.P. Yang, Phys. Rev. **150**, 321 (1966);*ibid.* **150**, 327 (1966).
[48] S.-J. Gu, H.-Q. Lin, and Y.-Q. Li, Phys. Rev. A **68**, 042330 (2003).
[49] O.F. Syljuasen, Phys. Rev. A **68**, 060301(R) (2003).
[50] J.-J. García-Ripoll and J.I. Cirac, New J. Phys. **5**, 76 (2003); L.-M. Duan, E. Demler, and M.D. Lukin, Phys. Rev. Lett. **91**, 090402 (2003).
[51] F. Verstraete, J.-J. García-Ripoll, and J.I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).
[52] J.-J. García-Ripoll, M. A. Martin-Delgado, J. I. Cirac, Phys. Rev. Lett. **93**, 250405 (2004).

[53] F.D.M. Haldane, Phys. Lett. **93A**, 464 (1983); Phys. Rev. Lett. **50**, 1153 (1983).
[54] M. den Nijs and K. Rommelse, Phys. Rev. B **40**, 4709 (1989); H. Tasaki, Phys. Rev. Lett. **66**, 798 (1991); T. Kennedy and H. Tasaki, Phys. Rev. B **45**, 304 (1992); M. Oshikawa, J. Phys. Condens. Matter **4**, 7469 (1992); H.V. Kruis, I.P. McCulloch, Z. Nussinov, J. Zaanen, Phys. Rev. B **70** 075109 (2004).
[55] We note that the degeneracy in the case of OBC could be lifted by placing $S = 1/2$ spins at the ends. For technical reasons we do not consider this situation.
[56] T. Kennedy, J. Phys.: Condens. Matter 2, 5737 (1990).
[57] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).
[58] F. Verstraete, J.I. Cirac, cond-mat/0407066.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

State Transfer in Permanently Coupled Quantum Chains

Daniel Burgarth [a], Vittorio Giovannetti [b], Sougato Bose [a]
[a] *Department of Physics & Astronomy, University College London, Gower St., London WC1E 6BT, UK.*
[b] *NEST-INFM & Scuola Normale Superiore, piazza dei Cavalieri 7, I-56126 Pisa, Italy.*

Abstract. We present a communication protocol for chains of permanently coupled qubits which achieves perfect quantum state transfer and which is efficient with respect to the number chains employed in the scheme. The system consists of M uncoupled identical quantum chains. Local control (gates, measurements) is only allowed at the sending/receiving end of the chains. Under a quite general hypothesis on the interaction Hamiltonian of the qubits a theorem is proved which shows that the receiver is able to asymptotically recover the messages by repetitive monitoring of her/his qubits. We show how two parallel Heisenberg spin chains can be used as quantum wires. Perfect state transfer with a probability of failure lower than P in a Heisenberg chain of N spin-1/2 particles can be achieved in a timescale of the order of $\frac{0.33\hbar}{J} N^{1.7} |\ln P|$. We demonstrate that our scheme is more robust to decoherence and non-optimal timing than any scheme using single spin chains. Arbitrarily perfect state transfer is even possible in the presence of random fluctuations in the couplings of a quantum chain. The scheme performs well for both spatially correlated and uncorrelated fluctuations if they are relatively weak (say 5%).

1. Introduction

Recently, much interest has been devoted to quantum communication with quantum chains [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18]. The main spirit of these articles is that permanently coupled systems can be used for the transfer of quantum information with *minimal control,* that is, only the sending and the receiving parties can apply gates to the system, but the part of the chain interconnecting them cannot be controlled during the communication process. A scheme with less control is obviously impossible. The first proposals [1,2] considered a regular spin chain with Heisenberg interactions. A physical implementation of this scheme was discussed in [3], and its channel capacity was derived in [4]. Already in [1,2] it was realized that such a transfer will in general not be perfect. The reason for the imperfect transfer is the *dispersion* of the information along the chain. This becomes worse as the chains get longer.

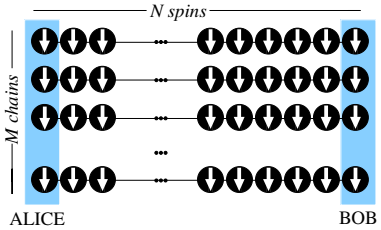


Figure 1. Schematic of the system: Alice and Bob operate M chains, each containing N spins. The spins belonging to the same chain interact through the Hamiltonian H which accounts for the transmission of the signal in the system. Spins of different chains do not interact. Alice encodes the information in the first spins of the chains by applying unitary transformations to her qubits. Bob recovers the message in the last spins of the chains by performing joint measurements.

Since then, many very interesting methods have been developed to improve the fidelity of the transfer. One method is to use Hamiltonians with engineered couplings [5,6,7,8,9,10,11,12] such that the dispersed information will "refocus" at the receiving end of the chain. Another approach is to encode and decode the information using multiple spins [13,14] to form Gaussian wave packets (which have a lower dispersion). This has been generalized [14] in an elegant way by using "phantom" spins such that a multiple-spin encoding can be achieved by only controlling two sending and receiving qubits. By using gapped systems [18, 15,16,17,19], the intermediate spins are only virtually excited, and the transfer has a very high fidelity.

In this article, we review another approach to state transfer based on a "dual-rail" encoding using *parallel* quantum channels [20,21,22,23,24,25]. We show that such a protocol achieves arbitrarily perfect transfer for nearly any type of quantum chain, transforming a heavily dispersive dynamic into one that can be used for state transfer. In this scheme, not only the control needed during the transfer is minimized (no local access along the chains is needed), but also the control needed to *design* the system in the first place. As the most extreme case, it can also be applied to *randomly* coupled spin chains.

The paper is organized as follows. In Sec. 2, we introduce the most general encoding using multiple parallel spin chains. The main convergence theorem is proved in Sec. 3. Finally in Sec. 4 we discuss the important special case where *two* spin chains are used. We show that this encoding is more stable against decoherence than other schemes and that it can be applied to randomly coupled chains.

2. The Model

Consider a linear chain of N spins interacting through the Hamiltonian H. For $n = 1, \cdots N$, we define the single excitation vector

$$|\boldsymbol{n}\rangle \equiv |00\cdots010\cdots0\rangle \,, \tag{1}$$

as the state of the chain in which the n-th spin is in the computational base vector $|1\rangle$ and the remaining $N-1$ qubits are in the state $|0\rangle$. Analogously we define $|\mathbf{0}\rangle \equiv |00\cdots0\rangle$ to be the state where all spins are in $|0\rangle$. We assume that $|\mathbf{0}\rangle$ is an eigenvector of H and that the N-dimensional subspace generated by the states $|\boldsymbol{n}\rangle$ is invariant under the time evolution $u(t) \equiv e^{-iHt/\hbar}$, i.e.

$$|\boldsymbol{n}\rangle \longrightarrow u(t)|\boldsymbol{n}\rangle = \sum_{n'=1}^{N} f_{n',n}(t)|\boldsymbol{n'}\rangle \ , \tag{2}$$

where $f_{n',n}(t) \equiv \langle \boldsymbol{n'}|e^{-iHt/\hbar}|\boldsymbol{n}\rangle$ is the probability amplitude that the excitation $|\boldsymbol{n}\rangle$ moves to $|\boldsymbol{n'}\rangle$ in the time interval t. A sufficient criterion for Eq. (2) is that H commutes with the z component of the total spin. A typical example is provided by a linear array of spins with Heisenberg interaction.

Assume now that the two communicating parties operate on M independent (i.e. non interacting) copies of the chain [27]. The idea is to use these copies to improve the overall fidelity of the communication. We assume Alice and Bob to control respectively the first and last qubit of each chain (see Fig. 1). By preparing any superposition of her spins Alice can in principle transfer up to M logical qubits. However, in order to improve the communication fidelity the two parties will find it more convenient to redundantly encode only a small number (say $Q(M) \leqslant M$) of logical qubits in the M spins. By adopting these strategies Alice and Bob are effectively sacrificing the efficiency $R(M) = Q(M)/M$ of their communication line in order to increase its fidelity. This is typical of any communication scheme and it is analogous to what happens in quantum error correction theory, where a single logical qubit is stored in many physical qubits. In the case of quantum chains this idea was originally discussed in Ref. [20] by showing that for $M = 2$ it is possible to achieve perfect state transfer of a single logical qubit with an efficiency equal to $1/2$. Here we will generalize such result by proving that there exist an optimal encoding-decoding strategy which asymptotically allows to achieve perfect state transfer *and* optimal efficiency, i.e.

$$\lim_{M\to\infty} R(M) = 1 \ . \tag{3}$$

Our strategy requires Alice to prepare superpositions of the M chains where $\sim M/2$ of them have a single excitation in the first location while the remaining are in $|\mathbf{0}\rangle$. Since in the limit $M >> 1$ the number of qubit transmitted is $\log \binom{M}{M/2} \approx M$, this architecture guarantees optimal efficiency (3). On the other hand, our protocol requires Bob to perform collective measurements on his spins to determine if all the $\sim M/2$ excitations Alice is transmitting arrived at his location. We will prove that by repeating these detections many times, Bob is able to recover the messages with asymptotically perfect fidelity.

2.1. Notation

Before beginning the analysis let us introduce some notation. In order to distinguish the M different chains we introduce the label $m = 1, \cdots, M$: in this formalism $|\boldsymbol{n}\rangle_m$ represents the state (1) of m-th chain with a single excitation

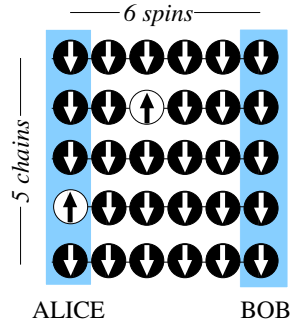


Figure 2. Example of our notation for $M = 5$ chains of length $N = 6$ with $K = 2$ excitations. The state above, given by $|\mathbf{0}\rangle_1 \otimes |\mathbf{3}\rangle_2 \otimes |\mathbf{0}\rangle_3 \otimes |\mathbf{1}\rangle_4 \otimes |\mathbf{0}\rangle_5$, has excitations in the chains $m_1 = 2$ and $m_2 = 4$ at the horizontal position $n_1 = 3$ and $n_2 = 1$. It is in the Hilbert space $\mathcal{H}(S_6)$ corresponding to the subset $S_6 = \{2, 4\}$ (assuming that the sets S_ℓ are ordered in a canonical way, i.e. $S_1 = \{1, 2\}$, $S_2 = \{1, 3\}$ and so on) and will be written as $|(3, 1); 6\rangle\rangle$. There are $\binom{5}{2} = 10$ different sets S_ℓ and the number of qubits one can transfer using these states is $\log_2 10 \approx 3$. The efficiency is thus given by $R \approx 3/5$.

in the n-th spin. In the following we will be interested in those configurations of the whole system where K chains posses a single excitation while the remaining $M - K$ are in $|\mathbf{0}\rangle$, as in the case

$$|\mathbf{1}\rangle_1 \otimes |\mathbf{1}\rangle_2 \cdots \otimes |\mathbf{1}\rangle_K \otimes |\mathbf{0}\rangle_{K+1} \cdots \otimes |\mathbf{0}\rangle_M \tag{4}$$

where for instance the first K chains have an excitation in the first chain location. Another more general example is given in Fig. 2. The complete characterization of these vectors is obtained by specifying *i) which* chains possess a single excitation and *ii) where* these excitations are located horizontally along the chains. In answering to the point *i)* we introduce the K-element subsets S_ℓ, composed by the labels of those chains that contain an excitation. Each of these subsets S_ℓ corresponds to a subspace of the Hilbert space $\mathcal{H}(S_\ell)$ with a dimension N^K. The total number of such subsets is equal to the binomial coefficient $\binom{M}{K}$, which counts the number of possibilities in which K objects (excitations) can be distributed amongst M parties (parallel chains). In particular for any $\ell = 1, \cdots, \binom{M}{K}$ the ℓ-th subset S_ℓ will be specified by assigning its K elements, i.e. $S_\ell \equiv \{m_1^{(\ell)}, \cdots, m_K^{(\ell)}\}$ with $m_j^{(\ell)} \in \{1, \cdots, M\}$ and $m_j^{(\ell)} < m_{j+1}^{(\ell)}$ for all $j = 1, \cdots, K$. To characterize the location of the excitations, point *ii)*, we will introduce instead the K-dimensional vectors $\vec{n} \equiv (n_1, \cdots, n_K)$ where $n_j \in \{1, \cdots, N\}$. We can then define

$$|\vec{\boldsymbol{n}}; \ell\rangle\rangle \equiv \bigotimes_{j=1}^{K} |\boldsymbol{n_j}\rangle_{m_j^{(\ell)}} \bigotimes_{m' \in \overline{S}_\ell} |\mathbf{0}\rangle_{m'} , \tag{5}$$

where $\overline{S}_\ell$ is the complementary of S_ℓ to the whole set of chains.

The state (5) represents a configuration where the j-th chain of the subset S_ℓ is in $|\boldsymbol{n_j}\rangle$ while the chains that do not belong to S_ℓ are in $|\mathbf{0}\rangle$ (see Fig. 2 for an explicit example). The kets $|\vec{\boldsymbol{n}}; \ell\rangle\rangle$ are a natural generalization of the states

$|\boldsymbol{n}\rangle_1 \otimes |\boldsymbol{0}\rangle_2$ and $|\boldsymbol{0}\rangle_1 \otimes |\boldsymbol{n}\rangle_2$ which were used for the "dual-rail encoding" in [20]. They are useful for our purposes because they are mutually orthogonal, i.e.

$$\langle\langle \vec{\boldsymbol{n}}; \ell | \vec{\boldsymbol{n}}'; \ell' \rangle\rangle = \delta_{\ell\ell'}\, \delta_{\vec{n}\vec{n}'} \ , \tag{6}$$

and their time evolution under the Hamiltonian does not depend on ℓ (cf. Eq. (12)). Among the vectors (5) those where all the K excitations are located at the beginning of the S_ℓ chains play an important role in our analysis. Here $\vec{n} = \vec{1} \equiv (1, \cdots, 1)$ and we can write

$$|\vec{\mathbf{1}}; \ell\rangle\rangle \equiv \bigotimes_{m \in S_\ell} |\mathbf{1}\rangle_m \bigotimes_{m' \in \overline{S}_\ell} |\mathbf{0}\rangle_{m'} \ . \tag{7}$$

According to Eq. (6), for $\ell = 1, \cdots, \binom{M}{K}$ these states form orthonormal set of $\binom{M}{K}$ elements. Analogously by choosing $\vec{n} = \vec{N} \equiv (N, \cdots, N)$ we obtain the orthonormal set of $\binom{M}{K}$ vectors

$$|\vec{\boldsymbol{N}}; \ell\rangle\rangle \equiv \bigotimes_{m \in S_\ell} |\boldsymbol{N}\rangle_m \bigotimes_{m' \in \overline{S}_\ell} |\mathbf{0}\rangle_{m'}, \tag{8}$$

where all the K excitations are located at the end of the chains.

2.2. Efficient Encoding

If all the M chains of the system are originally in $|\mathbf{0}\rangle$, the vectors (7) can be prepared by Alice by locally operating on her spins. Moreover since these vectors span a $\binom{M}{K}$ dimensional subspace, Alice can encode in the chain $Q(M, K) = \log_2 \binom{M}{K}$ qubits of logical information by preparing the superpositions,

$$|\Phi\rangle\rangle = \sum_\ell A_\ell \, |\vec{\mathbf{1}}; \ell\rangle\rangle \ , \tag{9}$$

with A_ℓ complex coefficients. The efficiency of such encoding is hence $R(M, K) = \frac{\log_2 \binom{M}{K}}{M}$ which maximized with respect to K gives,

$$R(M) = \frac{1}{M} \begin{cases} \log_2 \binom{M}{M/2} & \text{for } M \text{ even} \\ \log_2 \binom{M}{(M-1)/2} & \text{for } M \text{ odd} \ . \end{cases} \tag{10}$$

The Stirling approximation can then be used to prove that this encoding is asymptotically efficient (3) in the limit of large M, e.g.

$$\log_2 \binom{M}{M/2} \approx \log_2 \frac{M^M}{(M/2)^M} = M. \tag{11}$$

In the next section we show that the encoding (9) provides perfect state transfer by allowing Bob to perform joint measurements at its end of the chains.

2.3. Perfect Transfer

Since the M chains do not interact with each other and possess the same free Hamiltonian H, the unitary evolution of the whole system is described by $U(t) \equiv \otimes_m u_m(t)$, with $u_m(t)$ being the operator (2) acting on the m-th chain. The time evolved of the input $|\vec{\mathbf{1}}; \ell\rangle\!\rangle$ of Eq. (7) is thus equal to

$$U(t)|\vec{\mathbf{1}}; \ell\rangle\!\rangle = \sum_{\vec{n}} F[\vec{n}, \vec{1}; t] \; |\vec{\boldsymbol{n}}; \ell\rangle\!\rangle \; , \tag{12}$$

where the sum is performed for all $n_j = 1, \cdots, N$ and

$$F[\vec{n}, \vec{n'}; t] \equiv f_{n_1, n'_1}(t) \cdots f_{n_K, n'_K}(t) \; , \tag{13}$$

is a quantity which does *not* depend on ℓ. In Eq. (12) the term $\vec{n} = \vec{N}$ corresponds to having all the K excitations in the last locations of the chains. We can thus write

$$U(t)|\vec{\mathbf{1}}; \ell\rangle\!\rangle = \gamma_1(t)|\vec{\boldsymbol{N}}; \ell\rangle\!\rangle + \sqrt{1 - |\gamma_1(t)|^2} \; |\boldsymbol{\xi}(t); \ell\rangle\!\rangle \; , \tag{14}$$

where

$$\gamma_1(t) \equiv \langle\!\langle \vec{\boldsymbol{N}}; \ell | U(t) | \vec{\mathbf{1}}; \ell\rangle\!\rangle = F[\vec{N}, \vec{1}; t] \tag{15}$$

is the probability amplitude that all the K excitation of $|\vec{\mathbf{1}}; \ell\rangle\!\rangle$ arrive at the end of the chains, and

$$|\boldsymbol{\xi}(t); \ell\rangle\!\rangle \equiv \sum_{\vec{n} \neq \vec{N}} F_1[\vec{n}, \vec{1}; t] \; |\vec{\boldsymbol{n}}; \ell\rangle\!\rangle \; , \tag{16}$$

with

$$F_1[\vec{n}, \vec{1}; t] \equiv \frac{F[\vec{n}, \vec{1}; t]}{\sqrt{1 - |\gamma_1(t)|^2}}, \tag{17}$$

is a superposition of terms where the number of excitations arrived to the end of the communication line is strictly less then K. It is worth noticing that Eq. (6) yields the following relations,

$$\langle\!\langle \vec{\boldsymbol{N}}; \ell | \boldsymbol{\xi}(t); \ell' \rangle\!\rangle = 0, \quad \langle\!\langle \boldsymbol{\xi}(t); \ell | \boldsymbol{\xi}(t); \ell' \rangle\!\rangle = \delta_{\ell \ell'} \; , \tag{18}$$

which shows that $\{||\boldsymbol{\xi}(t); \ell\rangle\!\rangle\}$ is an orthonormal set of vectors which spans a subspace orthogonal to the states $|\vec{\boldsymbol{N}}; \ell\rangle\!\rangle$. The time evolution of the input state (9) follows by linearity from Eq. (14), i.e.

$$|\Phi(t)\rangle\!\rangle = \gamma_1(t) \; |\Psi\rangle\!\rangle + \sqrt{1 - |\gamma_1(t)|^2} \; |\overline{\Psi}(t)\rangle\!\rangle \; , \tag{19}$$

with

$$|\overline{\Psi}(t)\rangle\!\rangle \equiv \sum_{\ell} A_\ell \, |\boldsymbol{\xi}(t);\ell\rangle\!\rangle \ ,$$

$$|\Psi\rangle\!\rangle \equiv \sum_{\ell} A_\ell \, |\vec{\boldsymbol{N}};\ell\rangle\!\rangle \ . \tag{20}$$

The vectors $|\Psi\rangle\!\rangle$ and $|\overline{\Psi}(t)\rangle\!\rangle$ are unitary transformations of the input message (9) where the orthonormal set $\{|\vec{\mathbf{1}};\ell\rangle\!\rangle\}$ has been rotated into $\{|\vec{\boldsymbol{N}};\ell\rangle\!\rangle\}$ and $\{|\boldsymbol{\xi}(t);\ell\rangle\!\rangle\}$ respectively. Moreover $|\Psi\rangle\!\rangle$ is the configuration we need to have for perfect state transfer at the end of the chain. In fact it is obtained from the input message (9) by replacing the components $|\mathbf{1}\rangle$ (excitation in the first spin) with $|\boldsymbol{N}\rangle$ (excitation in the last spin). From Eq. (18) we know that $|\Psi\rangle\!\rangle$ and $|\overline{\Psi}(t)\rangle\!\rangle$ are orthogonal. This property helps Bob to recover the message $|\Psi\rangle\!\rangle$ from $|\Phi(t)\rangle\!\rangle$: he needs only to perform a collective measurement on the M spins he is controlling to establish if there are K or less excitations in those locations. The above is clearly a projective measure that can be performed without destroying the quantum coherence associated with the coefficients A_ℓ. Formally this can described by introducing the observable

$$\Theta \equiv \mathbb{1} - \sum_{\ell} |\vec{\boldsymbol{N}};\ell\rangle\!\rangle\langle\!\langle\vec{\boldsymbol{N}};\ell| \ . \tag{21}$$

A single measure of Θ on $|\Phi(t_1)\rangle\!\rangle$ yields the outcome 0 with probability $p_1 \equiv |\gamma_1(t_1)|^2$, and the outcome $+1$ with probability $1-p_1$. In the first case the system will be projected in $|\Psi\rangle\!\rangle$ and Bob will get the message. In the second case instead the state of the system will become $|\overline{\Psi}(t_1)\rangle\!\rangle$. Already at this stage the two communicating parties have a success probability equal to p_1. Moreover, as in [20], the channels have been transformed into a quantum erasure channel [29] where the receiver knows if the transfer was successful.

Consider now what happens when Bob fails to get the right answer from the measure. The state on which the chains is projected is explicitly given by

$$|\overline{\Psi}(t_1)\rangle\!\rangle = \sum_{\vec{n}\neq\vec{N}} F_1[\vec{n},\vec{1};t_1] \sum_{\ell} A_\ell |\vec{\boldsymbol{n}};\ell\rangle\!\rangle \ . \tag{22}$$

Let us now consider the evolution of this state for another time interval t_2. By repeating the same analysis given above we obtain an expression similar to (19), i.e.

$$|\Phi(t_2,t_1)\rangle\!\rangle = \gamma_2 \, |\Psi\rangle\!\rangle + \sqrt{1-|\gamma_2|^2} \, |\overline{\Psi}(t_2,t_1)\rangle\!\rangle \ , \tag{23}$$

where now the probability amplitude of getting all excitation in the N-th locations is described by

$$\gamma_2 \equiv \sum_{\vec{n}\neq\vec{N}} F[\vec{N},\vec{n};t_2] \ F_1[\vec{n},\vec{1};t_1].$$

In this case $|\overline{\Psi}(t)\rangle\!\rangle$ is replaced by

$$|\overline{\Psi}(t_2,t_1)\rangle\!\rangle = \sum_{\ell} A_\ell \, |\boldsymbol{\xi}(t_2,t_1);\ell\rangle\!\rangle \ , \tag{24}$$

with

$$|\boldsymbol{\xi}(t_2,t_1);\ell\rangle\rangle = \sum_{\vec{n}\neq\vec{N}} F_2[\vec{n},\vec{1};t_2,t_1]|\vec{\boldsymbol{n}};\ell\rangle\rangle,$$

and F_2 defined as in Eq. (26) (see below). In other words the state $|\Phi(t_2,t_1)\rangle\rangle$ can be obtained from Eq. (19) by replacing γ_1 and F_1 with γ_2 and F_2. Bob can hence try to use the same strategy he used at time t_1: i.e. he will check whether or not his M qubits contain K excitations. With (conditional) probability $p_2 \equiv |\gamma_2|^2$ he will get a positive answer and his quantum register will be projected in the state $|\Psi\rangle\rangle$ of Eq. (20). Otherwise he will let the system evolve for another time interval t_3 and repeat the protocol. Reiterating the above analysis it is possible to give a recursive expression for the conditional probability of success $p_q \equiv |\gamma_q|^2$ after $q-1$ successive unsuccessful steps. The quantity γ_q is the analogous of γ_2 and γ_1 of Eqs. (15) and (23). It is given by

$$\gamma_q \equiv \sum_{\vec{n}\neq\vec{N}} F[\vec{N},\vec{n};t_q]\ F_{q-1}[\vec{n},\vec{1},t_{q-1},\cdots,t_1]\ , \tag{25}$$

where

$$\begin{aligned} &F_{q-1}[\vec{n},\vec{1};t_{q-1},\cdots,t_1] \\ &\equiv \sum_{\vec{n}'\neq\vec{N}} \frac{F[\vec{N},\vec{n}';t_{q-1}]}{\sqrt{1-|\gamma_{q-1}|^2}} F_{q-2}[\vec{n}',\vec{1};t_{q-2},\cdots,t_1] \end{aligned} \tag{26}$$

and $F_1[\vec{n},\vec{1},t]$ is given by Eq. (17). In these equations $t_q,\cdots,t_1$ are the *times intervals* that occurred between the various protocol steps. Analogously the conditional probability of failure at the step q is equal to $1-p_q$. The probability of having $j-1$ failures and a success at the step j-th can thus be expressed as

$$\pi(j) = p_j(1-p_{j-1})(1-p_{j-2})\cdots(1-p_1)\ , \tag{27}$$

while the total probability of success after q steps is obtained by the sum of $\pi(j)$ for all $j=1,\cdots,q$, i.e.

$$P_q = \sum_{j=1}^{q} \pi(j)\ . \tag{28}$$

Since $p_j \geqslant 0$, Eq. (28) is a monotonic function of q. As a matter of fact in the next section we prove that under a very general hypothesis on the system Hamiltonian, the probability of success P_q converges to 1 in the limit of $q\to\infty$. This means that by repeating many times the collective measure described by Θ Bob is guaranteed to get, sooner or later, the answer 0 and hence the message Alice sent to him. In other words our protocol allows perfect state transfer in the limit of repetitive collective measures. Notice that the above analysis applies for all classes of subsets S_ℓ. The only difference between different choices of K is in the velocity of the convergence of $P_q \to 1$. In any case, by choosing $K \sim M/2$ Alice and Bob can achieve perfect fidelity *and* optimal efficiency.

3. Convergence Theorem

Here we show that if there exists no eigenvector $|e_m\rangle$ of the quantum chain Hamiltonian H which is orthogonal to $|\boldsymbol{N}\rangle$, than there is a choice of the times intervals $t_q, t_{q-1}, \cdots, t_1$ such P_q of Eq. (28) converges to 1 in the limit of $q \to \infty$. For the special case $M = 2$ and $K = 1$ this was numerically shown in Ref. [20].

The state of the system at a time interval of t_q after the $(q-1)$-th failure can be expressed in compact form as follows

$$|\Phi(t_q, \cdots, t_1)\rangle\!\rangle = \frac{U(t_q)\Theta U(t_{q-1})\Theta \cdots U(t_1)\Theta|\Phi\rangle\!\rangle}{\sqrt{(1-p_{q-1})\cdots(1-p_1)}}$$

with $U(t)$ the unitary time evolution generated by the system Hamiltonian, and with Θ the projection defined in Eq. (21). One can verify for instance that for $q = 2$, the above equation coincides with Eq. (23). [For $q = 1$ this is just (19) evaluated at time t_1]. By definition the conditional probability of success at step q-th is equal to

$$p_q \equiv |\langle\!\langle\Psi|\Phi(t_q, \cdots, t_1)\rangle\!\rangle|^2.$$

Therefore, Eq. (27) yields

$$\begin{aligned} \pi(q) &= |\langle\!\langle\Psi|U(t_q)\Theta U(t_{q-1})\Theta \cdots U(t_1)\Theta|\Phi\rangle\!\rangle|^2 \\ &= |\langle\!\langle\vec{\boldsymbol{N}};\ell|U(t_q)\Theta U(t_{q-1})\Theta \cdots U(t_1)\Theta|\vec{\mathbf{1}};\ell\rangle\!\rangle|^2 \ , \end{aligned} \tag{29}$$

where the second identity stems from the fact that, according to Eqs. (2) and (6), $U(t)\Theta$ preserves the orthogonality relation among states $|\vec{\boldsymbol{n}};\ell\rangle\!\rangle$ with distinct values of ℓ. Analogously to the cases of Eqs. (13) and (15), the second identity of (29) establishes that $\pi(q)$ can be computed by considering the transfer of the input $|\vec{\mathbf{1}};\ell\rangle\!\rangle$ for *arbitrary* ℓ. The expression (29) can be further simplified by noticing that for a given ℓ the chains of the subset $\overline{S}_\ell$ contribute with a unitary factor to $\pi(q)$ and can be thus neglected (according to (7) they are prepared in $|\mathbf{0}\rangle$ and do not evolve under $U(t)\Theta$). Identify $|\vec{\mathbf{1}}\rangle\!\rangle_\ell$ and $|\vec{\boldsymbol{N}}\rangle\!\rangle_\ell$ with the components of $|\vec{\mathbf{1}};\ell\rangle\!\rangle$ and $|\vec{\boldsymbol{N}};\ell\rangle\!\rangle$ relative to the chains belonging to the subset S_ℓ. In this notation we can rewrite Eq. (29) as

$$\pi(q) = |_\ell\langle\!\langle\vec{\boldsymbol{N}}|U_\ell(t_q)\Theta_\ell \ \cdots U_\ell(t_1)\Theta_\ell|\vec{\mathbf{1}}\rangle\!\rangle_\ell|^2 \ , \tag{30}$$

where $\Theta_\ell = \mathbb{1}_\ell - |\vec{\boldsymbol{N}}\rangle\!\rangle_\ell\langle\!\langle\vec{\boldsymbol{N}}|$ and $U_\ell(t)$ is the unitary operator $\otimes_{m\in S_\ell} u_m(t)$ which describes the time evolution of the chains of S_ℓ.

To prove that there exist a suitable choices of t_j such that the series (28) converges to 1 it is sufficient to consider the case $t_j = \tau > 0$ for all $j = 1, \cdots, q$: this is equivalent to selecting decoding protocols with constant measuring intervals.

By introducing the operator $T_\ell \equiv U_\ell(\tau)\Theta_\ell$, Eq. (30) becomes thus

$$\begin{aligned} &\pi(q) = |_\ell\langle\!\langle\vec{\boldsymbol{N}}|\ (T_\ell)^q|\vec{\mathbf{1}}\rangle\!\rangle_\ell|^2 \\ &=_\ell\langle\!\langle\vec{\mathbf{1}}|(T_\ell^\dagger)^q|\vec{\boldsymbol{N}}\rangle\!\rangle_\ell\langle\!\langle\vec{\boldsymbol{N}}|\ (T_\ell)^q|\vec{\mathbf{1}}\rangle\!\rangle_\ell = w(q) - w(q+1) \ , \end{aligned} \tag{31}$$

where

$$w(j) \equiv_\ell \langle\!\langle \vec{\mathbf{1}}|(T_\ell^\dagger)^j\ (T_\ell)^j|\vec{\mathbf{1}}\rangle\!\rangle_\ell = \|(T_\ell)^j|\vec{\mathbf{1}}\rangle\!\rangle_\ell\|^2\ , \tag{32}$$

is the norm of the vector $(T_\ell)^j|\vec{\mathbf{1}}\rangle\!\rangle_\ell$. Substituting Eq. (31) in Eq. (28) yields

$$P_q = \sum_{j=1}^{q} [w(j) - w(j+1)] = 1 - w(q+1) \tag{33}$$

where the property $w(1) = {}_\ell\langle\!\langle \vec{\mathbf{1}}|\Theta_\ell|\vec{\mathbf{1}}\rangle\!\rangle_\ell = 1$ was employed. Proving the thesis is hence equivalent to prove that for $q \to \infty$ the succession $w(q)$ nullifies. This last relation can be studied using properties of power bounded matrices [30]. In fact, by introducing the norm of the operator $(T_\ell)^q$ we have,

$$w(q) = \|(T_\ell)^q|\vec{\mathbf{1}}\rangle\!\rangle_\ell\|^2 \leqslant \|(T_\ell)^q\|^2 \leqslant c\left(\frac{1+\rho(T_\ell)}{2}\right)^{2q} \tag{34}$$

where c is a positive constant which does not depend on q [31] and where $\rho(T_\ell)$ is the spectral radius of T_ℓ, i.e. the eigenvalue of T_ℓ with maximum absolute value (N.B. even when T_ℓ is not diagonalizable this is a well defined quantity). Equation (34) shows that $\rho(T_\ell) < 1$ is a sufficient condition for $w(q) \to 0$. In our case we note that, given any normalized eigenvector $|\lambda\rangle\!\rangle_\ell$ of T_ℓ with eigenvalue λ we have

$$|\lambda| = \|T_\ell|\lambda\rangle\!\rangle_\ell\| = \|\Theta_\ell|\lambda\rangle\!\rangle_\ell\| \leqslant 1\ , \tag{35}$$

where the inequality follows from the fact that Θ_ℓ is a projector. Notice that in Eq. (35) the identity holds only if $|\lambda\rangle\!\rangle$ is also an eigenvector of Θ_ℓ with eigenvalue $+1$, i.e. only if $|\lambda\rangle\!\rangle_\ell$ is orthogonal to $|\vec{\boldsymbol{N}}\rangle\!\rangle_\ell$. By definition $|\lambda\rangle\!\rangle_\ell$ is eigenvector $T_\ell = U_\ell(\tau)\Theta_\ell$: therefore the only possibility to have the equality in Eq. (35) is that *i)* $|\lambda\rangle\!\rangle_\ell$ is an eigenvector of $U_\ell(\tau)$ (i.e. an eigenvector of the Hamiltonian H_ℓ^{tot} of the chain subset S_ℓ) and *ii)* it is orthogonal to $|\vec{\boldsymbol{N}}\rangle\!\rangle_\ell$. By negating the above statement we get a sufficient condition for the thesis. Namely, if all the eigenvectors $|\vec{E}\rangle\!\rangle_\ell$ of H_ℓ^{tot} are not orthogonal to $|\vec{\boldsymbol{N}}\rangle\!\rangle_\ell$ than the absolute values of the eigenvalues λ of T_ℓ are strictly smaller than 1 which implies $\rho(T_\ell) < 1$ and hence the thesis. Since the S_ℓ channels are identical and do not interact, the eigenvectors $|\vec{E}\rangle\!\rangle_\ell \equiv \bigotimes_{m\in S_\ell}|e_m\rangle_m$ are tensor product of eigenvectors $|e_m\rangle$ of the single chain Hamiltonians H. Using the notation introduced in Eq. (1) the sufficient condition becomes

$${}_\ell\langle\!\langle \vec{E}|\vec{\boldsymbol{N}}\rangle\!\rangle_\ell = \prod_{m\in S_\ell} {}_m\langle \boldsymbol{N}|e_m\rangle_m \neq 0\ , \tag{36}$$

which can be satisfied only if $\langle \boldsymbol{N}|e_m\rangle \neq 0$ for all eigenvectors $|e_m\rangle$ of the single chain Hamiltonian H. QED. Note that this condition is satisfied for any nearest-neighbor chain with exchange terms [21]. While we have proven here that for equal time intervals the probability of success is converging to unity, in practice one may use *optimal* measuring time intervals for a faster transfer [20]. We also point out that timing errors may delay the transfer, but will not decrease its fidelity.

4. Arbitrary and Conclusive Transfer With Two Parallel Heisenberg Spin Chains

After describing the general setup, let us turn to one specific but practically important example and consider two parallel and uncoupled ferromagnetic spin chains, each described by the Hamiltonian

$$H^{(i)} = -J \sum \vec{\sigma}_n^{(i)} \cdot \vec{\sigma}_{n+1}^{(i)}. \tag{37}$$

This model is exactly solvable, and the transition amplitude is given explicitly in [1]. For two chains, the efficient encoding (9) simply becomes the well-known dual-rail encoding [32]

$$\alpha \left|\mathbf{0}\right\rangle^{(1)} \otimes \left|\mathbf{1}\right\rangle^{(2)} + \beta \left|\mathbf{1}\right\rangle^{(1)} \otimes \left|\mathbf{0}\right\rangle^{(2)}. \tag{38}$$

In the case of two chains, it is also easy to specify the gates that are needed to encode and decode the state. If the state that Alice wants to transfer is located at the first qubit of the first chain,

$$\left|\psi\right\rangle_1^{(1)} \otimes \left|\mathbf{0}\right\rangle^{(2)} \equiv \left(\alpha \left|\mathbf{0}\right\rangle^{(1)} + \beta \left|\mathbf{1}\right\rangle^{(1)}\right) \otimes \left|\mathbf{0}\right\rangle^{(2)}, \tag{39}$$

she can put it in the state (38) by applying a NOT gate on the first qubit of system (2) controlled by the first qubit of system (1) being zero. This is assumed to take place in a much shorter timescale than the system dynamics. Even though a 2-qubit gate in solid state systems is difficult, such a gate for charge qubits has been reported [33]. For the same qubits, Josephson arrays have been proposed as single spin chains for quantum communication [3]. For this system, both requisites of our scheme are thus available[24]. In fact, the demand that Alice and Bob can do measurements and apply gates to their local qubits (i.e. the ends of the chains) will be naturally fulfilled in practice since we are suggesting a scheme to transfer information between quantum computers. The measurements [described by the observable (21)] that Bob needs to perform in order to check if the transfer was successful can be performed by applying a CNOT gate followed by a single-qubit measurement (see Fig. 3). In this case, the CNOT gate also takes care of decoding the state, such that after a measurement of success the state of the system will be

$$\left|\psi\right\rangle_N^{(1)} \otimes \left|\mathbf{0}\right\rangle^{(2)} \equiv \left(\alpha \left|\mathbf{0}\right\rangle^{(1)} + \beta \left|\mathbf{N}\right\rangle^{(1)}\right) \otimes \left|\mathbf{0}\right\rangle^{(2)}. \tag{40}$$

Our scheme is summarized in Fig. 3. As proven above, this setup allows to transfer one unknown qubit over an arbitrary distance with arbitrary high fidelity. The important question how long this transfer takes will be adressed in the next section.

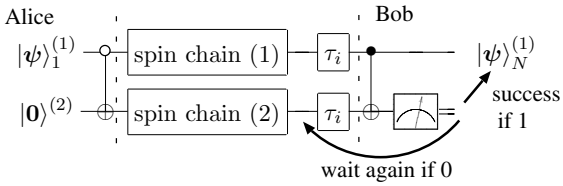


Figure 3. Quantum circuit representation of conclusive and arbitrarily perfect state transfer. The first gate at Alice's qubits represents a NOT gate applied to the second qubit controlled by the first qubit being zero. The qubits $|\psi\rangle_1^{(1)}$ on the left hand side represents an arbitrary input state at Alice's site, and the qubit $|\psi\rangle_N^{(1)}$ represents the same state, successfully transferred to Bob's site. The τ_i-gate represents the unitary evolution of the spin chains for a time interval of τ_i.

4.1. Estimation of the Timescale of the Transfer

The achievable fidelity is an important, but not the only criteria of a state transfer protocol. In this Section, we give an heuristic approach to estimate the time that it needs to achieve a certain fidelity in a Heisenberg spin chain. The comparison with numeric examples is confirming this approach.

Let us first describe the dynamic of the chain in a very qualitative way. Once Alice has initialized the system, an excitation wave packet will travel along the chain. As shown in [1], it will reach Bob at a time of the order of

$$\tau_{max} \approx \frac{\hbar N}{2J}, \tag{41}$$

with an amplitude of

$$|f_{N,1}(\tau_{max})|^2 \approx 1.35 N^{-2/3}. \tag{42}$$

It is then reflected and travels back and forth along the chain. Since the wave packet is also dispersing, it starts interfering with its tail, and after a couple of reflections the dynamic is becoming quite randomly. This effect becomes even stronger due to Bobs measurements, which change the dynamics by projecting away parts of the wave packet. However, $2\tau_{max}$ (the time it takes for a wave packet to travel twice along the chain) remains a good estimate of the timescale in which significant probability amplitude peaks at Bobs site occur, and Eq. (42) remains a good estimate of the amplitude of these peaks. Therefore, the joint probability of failure is expected to scale as

$$P(l) \approx \left(1 - 1.35 N^{-2/3}\right)^l \tag{43}$$

in a time of the order of

$$t(l) \approx 2\tau_{max} l = \frac{N\hbar l}{J}. \tag{44}$$

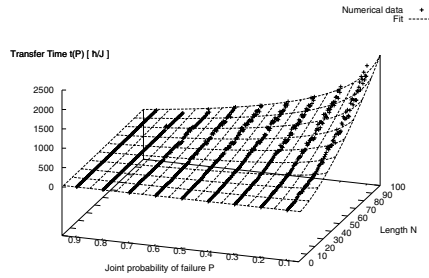


Figure 4. Time t needed to transfer a state with a given joint probability of failure P across a chain of length N. The points denote exact numerical data, and the fit is given by Eq. (46).

If we combine Eq. (43) and (44) and solve for the time $t(P)$ needed to reach a certain probability of failure P, we get

$$t(P) \approx \frac{0.51\hbar N^{5/3}}{J} \left|\ln P\right| . \tag{45}$$

We compare this rough estimate with exact numerical results in Fig. 4 . The best fit is given by

$$t(P) = \frac{0.33\hbar N^{5/3}}{J} \left|\ln P\right| . \tag{46}$$

We can conclude that the transmission time for arbitrarily perfect transfer is scaling not much worse with the length N of the chains than the single spin chain schemes. Despite of the logarithmic dependence on P, the time it takes to achieve high fidelity is still reasonable. For example, a system with $N = 100$ and $J = 20K * k_B$ will take approximately $1.3ns$ to achieve a fidelity of 99%. In many systems, decoherence is completely negligible within this timescale. For example, some Josephson junction systems [7] have a decoherence time of $T_\phi \approx 500ns$, while trapped ions have even larger decoherence times.

4.2. Decoherence and Imperfections

If the coupling between the spins J is very small, or the chains are very long, the transmission time may no longer be negligible with respect to the decoherence time (see Section 4.1). It is interesting to note that the dual-rail encoding then offers some significant general advantages over single chain schemes. Since we are suggesting a system-independent scheme, we will not study the effects of specific environments on our protocol, but just qualitatively point out its general advantages.

At least theoretically, it is always possible to cool the system down or to apply a strong magnetic field such that the environment is not causing further excitations. Then, there are two remaining types of quantum noise that will occur:

phase noise and amplitude damping. Phase noise is a serious problem and arises here *only* when an environment that can distinguish between spin flips on the first chain and spin flips on the second chain. It is therefore important that the environment cannot resolve their difference. In this case, the environment will only couple with the total z-component

$$S_{z,n} = \sigma_{z,n}^{(1)} + \sigma_{z,n}^{(2)} \tag{47}$$

of the spins of both chains at each position n. This has been discussed for spin-boson models in [34,35] but should also hold for spin environments as long as the chains are close enough. The qubit is encoded in a decoherence-free subspace [36] and the scheme is fully robust to phase noise. Even though this may not be true for all implementations of dual-rail encoding, it is worthwhile noticing it because such an opportunity does not exist *at all* for single chain schemes, where the coherence between two states with different total z-component of the spin has to be preserved. Having shown one way of avoiding phase noise, at least in some systems, we now proceed to amplitude damping.

The evolution of the system in presence of amplitude damping of a rate Γ can be easily derived using a quantum-jump approach [37]. Like for phase noise, it is necessary that the environment acts symmetrically on the chains. The dynamics is then given by an effective non-unitary Hamiltonian

$$H_{eff} = H + i\Gamma \sum_n \left(S_{z,n} + 2\right)/2 \tag{48}$$

if no jump occurs, and the effect of a jump is given by the operator

$$\sum_n S_n^- , \tag{49}$$

which will put the system in the ground state. As this can be solved analytically, we do not go into numerics. The state of the system before the first measurement conditioned on no jump is given by

$$e^{-\Gamma t} \sum_{n=1}^{N} f_{n,1}(t) \left|\mathbf{s}(n)\right\rangle , \tag{50}$$

and this happens with the probability of $e^{-2\Gamma t}$ (the norm of the above state). If a jump occurs, the system will be in the ground state

$$\sqrt{1 - e^{-2\Gamma t}} \left|\mathbf{0}\right\rangle^{(1)} \otimes \left|\mathbf{0}\right\rangle^{(2)} . \tag{51}$$

The density matrix at the time t is given by a mixture of (50) and (51). In case of (51), the quantum information is completely lost and Bob's error check qubit will never show success. If Bob however measures a success, it is clear that no jump has occurred and he has the perfectly transferred state. Therefore the protocol *remains conclusive*, but the success probability is lowered by $e^{-2\Gamma t}$. This result is

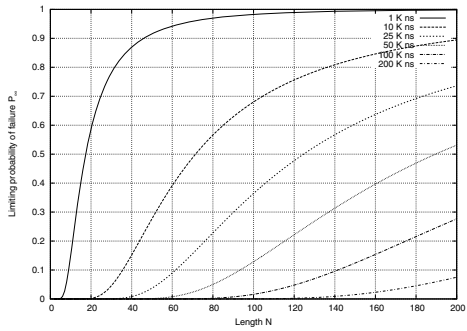


Figure 5. The minimal joint probability of failure $P(l)$ for chains with length N in the presence of amplitude damping. The parameter J/Γ of the curves is the coupling of the chain (in Kelvin) divided by the decay rate (ns^{-1}).

still valid for multiple measurements, which leave the state (51) unaltered. The probability of a successful transfer at each particular measurement l will decrease by $e^{-2\Gamma t(l)}$, where $t(l)$ is the time of the measurement. After a certain number of measurements, the *joint* probability of failure will no longer decrease. Thus the transfer will no longer be *arbitrarily* perfect, but can still reach a very high fidelity. Some numerical examples of the minimal joint probability of failure that can be achieved,

$$P_\infty \equiv \lim_{l\to\infty} P(l) \tag{52}$$

$$\approx \prod_{l=1}^{\infty} \left(1 - 1.35 N^{-2/3} e^{-\frac{2\Gamma N\hbar}{J} l}\right) \tag{53}$$

are given in Fig. 5. For $J/\Gamma = 50K\ ns$ nearly perfect transfer is still possible for chains up to a length of $N \approx 40$. In a single Heisenberg chain using the scheme described in [1], this system could only achieve a fidelity of 0.23 when transferring an exitation.

Even if the amplitude damping is not symmetric, its effect is weaker than in single spin schemes. This is because it can be split in a symmetric and asymmetric part. The symmetric part can be overcome with the above strategies. For example, if the amplitude damping on the chains is Γ_1 and Γ_2 with $\Gamma_1 > \Gamma_2$, the state (50) will be

$$\sum_{n=1}^{N} f_{n,1}(t) \left\{\alpha e^{-\Gamma_2 t} |\mathbf{0n}\rangle + \beta e^{-\Gamma_1 t} |\mathbf{n0}\rangle\right\} \tag{54}$$

$$= e^{-\Gamma_2 t} \sum_{n=1}^{N} f_{n,1}(t) \left\{\alpha |\mathbf{0n}\rangle + \beta e^{-(\Gamma_1-\Gamma_2)t} |\mathbf{n0}\rangle\right\} \tag{55}$$

$$\approx e^{-\Gamma_2 t} \sum_{n=1}^{N} f_{n,1}(t) |\mathbf{s}(n)\rangle \tag{56}$$

provided that $t \ll (\Gamma_1 - \Gamma_2)^{-1}$. Using a chain of length $N = 20$ with $J = 20K * k_B$ and $\Gamma_1^{-1} = 4ns, \Gamma_2^{-1} = 4.2ns$ we would have to fulfill $t \ll 164ns$. We could perform approximately 10 measurements (cf. Eq. (44)) without deviating too much from the state (56). In this time, we can use our protocol in the normal way. The resulting success probability given by the finite version of Eq. (53) would be 75%. A similar reasoning is valid for phase noise, where the environment can be split into common and seperate parts. If the chains are close, the common part will dominate and the seperate parts can be neglected for short times.

Finally, let us mention another advantage of our scheme. In single chain schemes, Bob has to extract the state precisely at an optimal time to obtain it with high fidelity. Our scheme is robust to the errors in this. Even if Bob measures to extract his state at an incorrect (non-optimal) time, he will receive the perfect state conditional on his measurement outcome. If he is unsuccessful, he simply tries again, without having Alice to resend. Also, due to the conclusive nature of the protocol, once Bob has received the state, the rest of the channel is automatically in the ground state. Therefore channel memory effects [38] are not important and the channel does not need to be reset for the next transfer (as opposed to many of the existing schemes [1,2,39,40]).

4.3. Randomly Coupled Chains

The main requirement for the above discussion is that two (or more) *identical* quantum chains have to be designed. While this is not so much a theoretical problem, for possible experimental realizations of the scheme [3] the question arises naturally how to cope with slight asymmetries of the channels. The purpose of this section is to demonstrate that in many cases, perfect state transfer with dual-rail encoding is possible for quantum chains with differing Hamiltonians. By doing so, we also offer a solution to another and perhaps more *general* problem: if one implements *any* of the state transfer scheme, the Hamiltonians will always be different from the theoretical ones by some random perturbation. This will lead to a decrease of fidelity in particular where specific energy levels were assumed (see [41] for an analysis of fluctuations affecting the scheme [5]). Also in general, random systems can lead to a Anderson localization [42] of the eigenstates (and therefore to low fidelity transport of quantum information). This problem can be avoided by using the above scheme, but restricting the measurement times to times where a certain condition (66) is met. In this section we will briefly show how this method works for the first measurement ("conclusive transfer"). A more detailed analysis can be found in [22].

Consider two parallel and uncoupled chains,e described by the two Hamiltonians $H^{(1)}$ and $H^{(2)}$ acting on the corresponding Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. The total Hamiltonian is thus

$$H = H^{(1)} \otimes I^{(2)} + I^{(1)} \otimes H^{(2)}, \tag{57}$$

and the time evolution operator factorizes as

$$U(t) = \exp(-iHt) \tag{58}$$

$$= \exp\left(-iH^{(1)}t\right) \otimes \exp\left(-iH^{(2)}t\right). \tag{59}$$

One important example in the context of quantum chains is given by

$$H^{(1)} = \sum_{n=1}^{N-1} J(1+\delta_n^{(1)})\vec{\sigma}_n^{(1)} \cdot \vec{\sigma}_{n+1}^{(1)}, \tag{60}$$

$$H^{(2)} = \sum_{n=1}^{N-1} J(1+\delta_n^{(2)})\vec{\sigma}_n^{(2)} \cdot \vec{\sigma}_{n+1}^{(2)}, \tag{61}$$

where $\delta_n^{(i)}$ are uniformly distributed (but time independent) random numbers from the interval $[-\Delta, \Delta]$. This corresponds to the case where a pair of unmodulated chains have been manufactured, but some random noise $\delta_n^{(i)}$ affects the couplings between the qubits. As above, Alice encodes the state as

$$\alpha \left|\mathbf{0},\mathbf{1}\right\rangle + \beta \left|\mathbf{1},\mathbf{0}\right\rangle. \tag{62}$$

This is a superposition of an excitation in the first qubit of the first chain and an excitation in the first qubit of the second chain. The state will evolve into

$$\sum_{n=1}^{N} \left\{\alpha g_{n,1}(t) \left|\mathbf{0},\mathbf{n}\right\rangle + \beta f_{n,1}(t) \left|\mathbf{n},\mathbf{0}\right\rangle\right\}. \tag{63}$$

There are now *two* different probability amplitudes given by

$$f_{n,1}(t) \equiv \left\langle \mathbf{n},\mathbf{0}\right| U(t) \left|\mathbf{1},\mathbf{0}\right\rangle \tag{64}$$

$$g_{n,1}(t) \equiv \left\langle \mathbf{0},\mathbf{n}\right| U(t) \left|\mathbf{0},\mathbf{1}\right\rangle. \tag{65}$$

We may, however, find a time t_1 such that the modulus of their amplitudes at the last spins are the same (see Fig. 6),

$$g_{N,1}(t_1) = e^{i\phi_1} f_{N,1}(t_1). \tag{66}$$

At this time, the state (63) can be written as

$$\sum_{n=1}^{N-1} \left\{\alpha g_{n,1}(t_1) \left|\mathbf{0},\mathbf{n}\right\rangle + \beta f_{n,1}(t_1) \left|\mathbf{n},\mathbf{0}\right\rangle\right\} + f_{N,1}(t_1) \left\{e^{i\phi_1}\alpha \left|\mathbf{0},\mathbf{N}\right\rangle + \beta \left|\mathbf{N},\mathbf{0}\right\rangle\right\}. \tag{67}$$

Bob decodes the state by applying a CNOT gate on his two qubits, with the first qubit as the control bit. The state thereafter is

$$\sum_{n=1}^{N-1} \left\{\alpha g_{n,1}(t_1) \left|\mathbf{0},\mathbf{n}\right\rangle + \beta f_{n,1}(t_1) \left|\mathbf{n},\mathbf{0}\right\rangle\right\} + f_{N,1}(t_1) \left\{e^{i\phi_1}\alpha \left|\mathbf{0}\right\rangle^{(1)} + \beta \left|\mathbf{N}\right\rangle^{(1)}\right\} \otimes \left|\mathbf{N}\right\rangle^{(2)}. \tag{68}$$

Bob then measures his second qubit. Depending on the outcome of this measurement, the systems will either be in the state

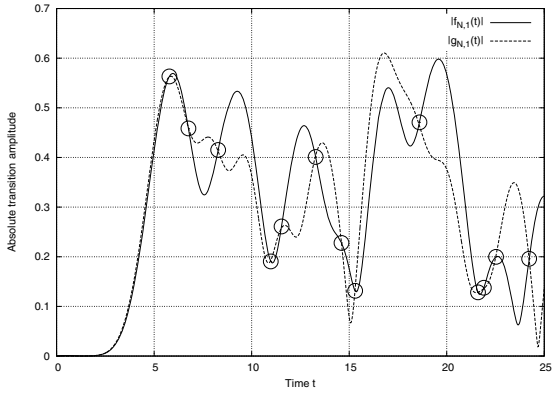


Figure 6. The absolute values of the transition amplitudes $f_{N,1}(t)$ and $g_{N,1}(t)$ for two Heisenberg chains of length $N = 10$. The couplings strengths of both chains were chosen randomly from the interval $[0.8J, 1.2J]$. The circles show times where Bob can perform measurements without gaining information on α and β.

$$\frac{1}{\sqrt{p_1}} \sum_{n=1}^{N-1} \{\alpha g_{n,1}(t_1) \left|\mathbf{0}, \mathbf{n}\right\rangle + \beta f_{n,1}(t_1) \left|\mathbf{n}, \mathbf{0}\right\rangle\} \tag{69}$$

or in

$$\left\{e^{i\phi_1} \alpha \left|\mathbf{0}\right\rangle^{(1)} + \beta \left|\mathbf{N}\right\rangle^{(1)}\right\} \otimes \left|\mathbf{N}\right\rangle^{(2)}, \tag{70}$$

where $p_1 = 1 - |f_{N,1}(t_1)|^2 = 1 - |g_{N,1}(t_1)|^2$ is the probability that Bob has *not* received the state. The state (70) corresponds to the correctly transferred state with a known phase error (which can be corrected by Bob using a simple phase gate). If Bob finds the system in the state (69), the transfer has been unsuccessful, but the information is still in the chain. We thus see that conclusive transfer is still possible with randomly coupled chains as long as the requirement (66) is met. This requirement is further discussed and generalized in [22], where we also give some numerical examples. It is shown numerically that the dual-rail scheme can still achieve arbitrarily perfect transfer for a uniformly coupled Heisenberg hamiltonian with random noise [Eq. (60) and (61)] on the coupling strengths (both for the case of spatially correlated and uncorrelated fluctuations) if the noise is relatively low. Moreover, for any two quantum chains, we show that Bob and Alice can check whether their system is capable of dual-rail transfer without directly measuring their Hamiltonians or local properties of the system along the chains but by only measuring *their* part of the system.

Acknowledgements

VG acknowledges the support of the European Community under contracts IST-SQUIBIT, IST-SQUBIT2, and RTN-Nanoscale Dynamics. DB acknowledges the

support of the UK Engineering and Physical Sciences Research Council, Grant Nr. GRS62796/01 and the QIPIRC.

References

[1] S. Bose, Phys. Rev. Lett. **91**, 207901 (2003).
[2] V. Subrahmanyam, Phys. Rev. A **69**, 034304 (2004).
[3] A. Romito, R. Fazio, C. Bruder, Phys. Rev. B **71**, 100501 (2005).
[4] V. Giovannetti and R. Fazio, Phys. Rev. A **71**, 032314 (2005).
[5] M. Christandl, N. Datta, A. Ekert and A. J. Landahl, Phys. Rev. Lett. **92**, 187902 (2004).
[6] C. Albanese, M. Christandl, N. Datta and A. Ekert, Phys. Rev. Lett. **93**, 230502 (2004).
[7] M. Christandl, N. Datta, T. C. Dorlas, A. Ekert, A. Kay, and A. J. Landahl, Phys. Rev. A **71**, 032312 (2005).
[8] P. Karbach and J. Stolze, Phys. Rev. A **72**, 030301 (R) (2005).
[9] G. M. Nikolopoulos, D. Petrosyan and P. Lambropoulos, J. Phys.: Condens. Matter **16**, 4991 (2004).
[10] G. M. Nikolopoulos, D. Petrosyan and P. Lambropoulos, Europhys. Lett. **65**, 297 (2004).
[11] MH Yung, DW Leung and S. Bose, Quant. Inf. & Comp. **4**, 174 (2004).
[12] MH Yung and S. Bose, Phys. Rev. A **71**, 032310 (2005).
[13] T. J. Osborne and N. Linden, Phys. Rev. A **69**, 052315 (2004).
[14] H. L. Haselgrove, eprint quant-ph/0404152.
[15] Y.Li, T.Shi, B.Chen, Z.Song, C.P.Sun, Phys. Rev. A **71**, 022301 (2005).
[16] T. Shi, Ying Li, Z. Song, C. P. Sun, Phys. Rev. A **71**, 032309 (2005).
[17] Z. Song, C. P. Sun, Fizika Nizkikh Temperatur **31**, Nos. 8/9 (2005).
[18] M. B. Plenio and F. L. Semiao, New. J. Phys. **7**, 73 (2005).
[19] A. Wojcik, T. Luczak, P. Kurzynski, A. Grudka, T. Gdala, M. Bednarska, Phys. Rev. A **72**, 034303 (2005).
[20] D. Burgarth and S. Bose, Phys. Rev. A **71**, 052315 (2005).
[21] D. Burgarth, S. Bose and V. Giovannetti, J. Phys. A: Math. Gen. **38** 6793 (2005).
[22] D. Burgarth and S. Bose, New J. Phys. **7**, 135 (2005).
[23] B. Vaucher, D. Burgarth and S. Bose, J. Opt. B **7**, 356 (2005)
[24] A. Lyakhov and C. Bruder, New J. Phys. **7**, 181 (2005).
[25] A. Kay and M. Ericsson, New J. Phys. 7, **143** (2005).
[26] C. H. Bennett and P. W. Shor, IEEE Trans. Inf. Theory **44**, 2724 (1998).
[27] This is quite a common attitude in quantum information theory [26] where successive uses of a memoryless channel are formally described by introducing many parallel copies of the channel (see [4] for a discussion on the possibility of applying this formal description to quantum chain models). Moreover for the case at hand the assumption of Alice and Bob dealing with "real" parallel chain seems reasonable also from a practical point of view [28].
[28] N. Motoyama, H. Eisaki, and S. Uchida, Phys. Rev. Lett. **76**, 3212 (1996); P. Gambardella, A. Dallmeyer, K. Maiti, M. C. Malagoli, W. Eberdardt, K. Kern, and C. Carbone, Nature **416**, 301 (2002).
[29] C. H. Bennett, D. P. DiVincenzo, and John A. Smolin, Phys. Rev. Lett **78**, 3217 (1997).
[30] J. R. Schott, Matrix Analysis for Statistics, Wiley-Interscience (1996).
[31] If S is the similarity transformation that puts T_ℓ into the Jordan canonical form, i.e. $J = S^{-1}T_\ell S$, then c is given explicitly by $c = \|S\| \, \|S^{-1}\|$.

[32] I. L. Chuang and Y. Yamamoto, Phys. Rev. Lett. **76**, 4281 (1996).
[33] T. Yamamoto, YA Pashkin, O. Astafiev, Y. Nakamura and JS Tsai, Nature **425**, 941 (2003).
[34] G.M. Palma, K.A. Suominen and A.K. Ekert, Proc. R. Soc. Lond. A **452**, 567 (1996).
[35] W.Y. Hwang, H. Lee, D.D. Ahn, S.W. Hwang, Phys. Rev. A **62**, 062305 (2000).
[36] A. Beige, D.Braun and PL. Knight, New Journal of Physics **2**, 22 (2000).
[37] M. Plenio and P. Knight, Rev. Mod. Phys. **70**, 101 (1998).
[38] D. Kretschmann, R. F. Werner, quant-ph/0502106.
[39] F. Verstraete, M. Popp and J. I. Cirac, Phys. Rev. Lett. **92**, 027901.
[40] F. Verstraete, M. A. Martin-Delgado and J. I. Cirac, Phys. Rev. Lett. **92**, 087201 (2004) .
[41] G. De Chiara, D. Rossini, S. Montangero, and R. Fazio, Phys. Rev. A **72**, 012323 (2005).
[42] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Geometric Effects in Spin Chains

Marie Ericsson [1], Alastair Kay
Department of Applied Mathematics and Theoretical Physics (DAMTP), University of Cambridge, UK

Abstract. An N qubit spin chain with engineered nearest neighbor couplings can perform perfect state transfer. The Hamiltonian of the system is analogous to an x-rotation of a large spin, i.e. a spin $J = \frac{1}{2}(N-1)$ particle rotating about the x-axis from its $|\, M_z = J\rangle$ to $|\, M_z = -J\rangle$ state. Here, we extend this analogy to consider rotations also around the z-axis and we show how to generate geometric effects, both Abelian and Non-Abelian, in spin chains.

Keywords. Spin chains, geometric phase

1. Introduction

It has recently been shown how a chain of spins can be coupled together to allow the perfect transfer of a quantum state between opposite ends of the chain [1,2,3,4,5,6]. The first of these, defined for an N qubit chain, was based on an analogy with a spin $J = \frac{1}{2}(N-1)$ particle rotating about the x-axis from its $|\, M_z = J\rangle$ to $|\, M_z = -J\rangle$ state. We extend the analogy to achieve rotations about the z-axes. We show how to perform a geometric phase gate on a state transported in a single spin chain and also how to apply a geometric Hadamard, i.e. a non-Abelian geometric phase gate, on a state that is transmitted through a spin network [7].

2. Geometric Effects in a Single Spin Chain

The Hamiltonian for perfect state transfer is given by [1,3,2]

$$H = \lambda \sum_{i=1}^{N-1} \frac{\sqrt{i(N-i)}}{4} (\sigma_x^i \sigma_x^{i+1} + \sigma_y^i \sigma_y^{i+1}) = \bigoplus_i H_i. \tag{1}$$

The total z-component of the spins commute with the Hamiltonian and is therefore conserved. This means that the ratio of spin up and spin down is the same through the evolution provided there is no external interaction. Taking the subspace where we only have one excitation in the system we can write the Hamiltonian, using basis states $|\, n\rangle$ to denote the presence of the excitation on qubit n, as

[1]Correspondence to: Marie Ericsson, Department of Applied Mathematics and Theoretical Physics (DAMTP), University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK. Tel.: +44 (0) 1223 760383; Fax: +44 (0) 1223 765 900 ; E-mail: M.Ericsson@damtp.cam.ac.uk

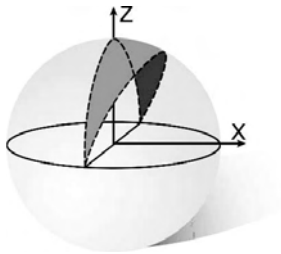


Figure 1. By performing rotations around great circles, there are no dynamical phases, making the result robust against some types of error.

$$H_1 = \frac{\lambda}{2}\begin{pmatrix} 0 & \sqrt{N-1} & 0 & \dots & 0 \\ \sqrt{N-1} & 0 & \sqrt{2(N-2)} & \dots & 0 \\ 0 & \sqrt{2(N-2)} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \sqrt{N-1} \\ 0 & 0 & 0 & \sqrt{N-1} & 0 \end{pmatrix}.$$

This is just the rotation matrix J_x for a particle of spin $J = \frac{1}{2}(N-1)$, scaled by a strength parameter λ. If the single excitation is on qubit n, then this is equivalent in the spin picture to being in the state $|\, M_z = J + 1 - n\rangle$. We can thus view state transfer between opposite ends of the chain as a rotation by an angle π around the x-axis of the Bloch sphere. Similarly, we can use the Bloch sphere picture to understand that if we rotate by an angle of $\pi/2$, we will move from the state $|\, M_z = J\rangle$ to $|\, M_y = J\rangle$. However, with only the Hamiltonian J_x, the rotations that we can generate are limited. We will thus introduce a magnetic field gradient over the spin chain. This is equivalent to a rotation around the z-axis of the spin-J particle, and takes the form

$$\frac{\kappa}{2}\begin{pmatrix} N-1 & 0 & 0 & \dots & 0 \\ 0 & N-3 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 1-N \end{pmatrix}$$

in the first excitation subspace, where κ is a coupling strength that we control.

The situation that we envisage is when we have a chain with a fixed Hamiltonian, eqn. (1), of known λ and that we can apply a gradient magnetic field by varying κ. This is sufficient to be able to implement a geometric rotation with the following protocol. Place a single excitation at one end of the chain i.e. $|\, M_z = J\rangle$. Then wait for a time $\pi/(2\lambda)$, resulting in $|\, M_y = J\rangle$. Switch on the magnetic gradient field at strength κ for a time $\pi/\sqrt{\lambda^2+\kappa^2}$, which rotates the state to $|\, M_y = -J\rangle$. Finally switch off the magnetic field, and wait for a time $\pi/(2\lambda)$. This protocol returns the single excitation to its starting point, having followed great circles of the Bloch sphere, as depicted in Fig. 1. This means that no dynamic phase is generated. However, a solid angle $\Omega = 2\tan^{-1}\left(\frac{\lambda}{\kappa}\right)$ has been carved out, and the geometric phase of the state is thus [8,9] $\gamma = \frac{1}{2}(N-1)\Omega$.

Hence, by selecting κ we apply a different geometric phsae gate to the state on the chain. Such geometric rotations have additional stability against errors in comparison to normal gates.

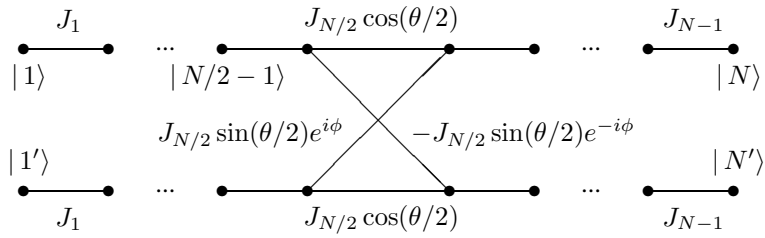


Figure 2. Two coupled perfect state transfer chains. Adiabatic varying of the parameters θ and ϕ gives a non-abelian geometric phase.

3. Non-Abelian Geometric Gates

We have seen that by introducing a limited form of control (a magnetic field gradient), a phase gate can be made geometric in nature, yielding a robustness against a variety of errors. We will now show how to create a Hadamard gate in the dual rail encoding, for which we invoke a non-abelian geometric phase.

To realise a non-abelian geometric phase, we require spin networks with degenerate eigenvalues. Consider, therefore, two perfect state transfer chains, both of length N. Our quantum state is encoded as

$$|0_L\rangle = |01\rangle$$
$$|1_L\rangle = |10\rangle$$

across the first qubit in each of the the two chains. There are degenerate eigenvalues because everything is repeated in comparison to a single chain. We can make some small modification to these chains, linking them as shown in Fig. 2. This allows the introduction of two free parameters, θ and ϕ, that control the coupling.

We will now consider a path in the parameter space (θ, ϕ) that change adiabatically from $(0,0) \to (\theta,0) \to (\theta,\phi) \to (0,\phi)$. If we start our state in a degenerate eigenspace, then it remains in this space when these parameters are varied adiabatically. In our case all degenerated pairs of eigenstates will be affected in the same way when varying θ and ϕ. To find the evolution of the state, we differentiate the degenerate eigenvectors with respect to θ and ϕ [10], and find that

$$\frac{\partial}{\partial\zeta}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = A_\zeta \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

for $\zeta \in \{\theta, \phi\}$, where a_i are the amplitudes of our state in the two different eigenvectors. The matrices, A_ζ, are given by

$$A_\theta = i\frac{1}{4}\sigma_y$$
$$A_\phi = \frac{i}{4}(\sin\theta\sigma_x + (1+\cos\theta)\sigma_z)$$

(up to an identity matrix, which only contributes a global phase).

Since both A_ζ are independent of ζ, evaluating the evolution just involves exponentiating the matrices. For convenience, we define the variable s such that $A_\phi^2 = -s^2\mathbf{1}$. We can therefore evaluate the whole evolution as we vary the parameters

$$U = e^{-A_\theta\theta}e^{A_\phi\phi}e^{A_\theta\theta} = \mathbf{1}\cos(s\phi) + \frac{i}{s}e^{-i\frac{\theta}{2}\sigma_y}\sin(s\phi)\,A_\phi.$$

The Hadamard gate is defined as $H = (X+Z)/\sqrt{2}$ and is the resulting evolution if we first select $\phi = \frac{\pi}{2s}$ for the coefficient in front of the identity to be zero and then set $\theta = \pi/4$ and therefore $\phi = \pi/\cos(\pi/8)$.

4. Conclusions

In this paper we have taken spin chains that can perform perfect state transfer and, with the addition of some extra control, we have shown how to implement Abelian and non-Abelian geometric gates on states of the chain. These have a robustness against a wide range of errors while allowing all possible single-qubit unitary operations.

Acknowledgements

Financial support from Swedish research council and traveling support from EU programmes RESQ and TopQip are acknowledge.

References

[1] M. Christandl et al. Phys. Rev. Lett. **92** (2004), 187902.
[2] C. Albanese et al. Phys. Rev. Lett. **93** (2004), 230502.
[3] M. Christandl et al. Phys. Rev. A **71** (2005), 032312.
[4] P. Karbach and J. Stolze, quant-ph/0501007.
[5] M.-H. Yung and S. Bose, Phys. Rev. A **71** (2005), 032310.
[6] T. Shi et al., Phys. Rev. A **71** (2005), 032309.
[7] A. Kay and M. Ericsson, New J. Phys. **7** (2005), 143.
[8] A. Ekert et al., J. Mod. Opt. **47** (2000), 2501.
[9] A. Shapere and F. Wilczek (ed), Geometric Phases in Physics (Advanced Series in Mathematical Physics vol 5, (1989) Singapore: World Scientific.
[10] V. Vedral., Int. J. Quantum Inf. **1** (2003), 1.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Quantum Walks and Decoherence on a 1+1 Lattice

Isabelle Herbauts

Queen Mary, University of London, Physics Department, London U.K.

Abstract. We present a flexible code to analyse particular models of quantum field theories on a lattice, allowing the investigation of discrete quantum field theories with varying amounts of "intervention". This intervention can be thought of either as the decohering effect of a local coupling to an environment or a fundamental new dynamics for the field, in which case the theory is an example of a causal, discrete collapse model.

Keywords. Wave function collapse, lightcone field theory, 1+1 lattice

1. Introduction

Collapse models (for a review, see [1]) have been developed to account for the dichotomy of standard quantum mechanics: while the Schrödinger equation is linear and deterministic, the wave packet reduction, invoked in the somewhat ill-defined occurrence of a "measurement" or "observation", is non-linear and stochastic. Collapse models postulate non-unitary "hits" (reductions of the wave function) superimposed on the Schrödinger equation as a universal dynamical process, independent of any intensional measurement. In the Continuous Spontaneous Localisation model (CSL), a random fluctuating term is superposed on the Schrödinger equation (continuous hitting), which causes collapse of the wave function in the appropriate situations.

We present a collapse model on a lattice and show that in a purely unitary version of this model, the probability distribution generated by the evolution can be interpreted as a quantum walk. Introduction of partial collapses at discrete time and places provide an interpretation in which the usual coins of a quantum walk are subjected to decoherence (partial collapse). While in the collapse models the hittings are postulated as a fundamental new dynamics for the field, in the standard no-collapse quantum theory, partial collapses can be ascribed to the decohering effect of a local coupling to an environment. Such decohering quantum walks have been studied, amongst others, in [2,3,4]. The simulations presented here are a good testing ground for investigations of the different interpretations, as well as allowing the generalisation multiple walkers.

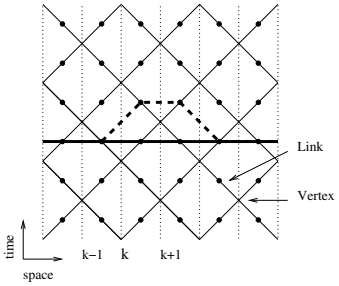


Figure 1. The ligtcone 1+1 lattice with discrete spacetime points on each links and transfer R-matrices at each vertex. A constant time surface is shown by the thick plain line, while the thick dotted line is obtained from it by an elementary motion, when only a single vertex is crossed.

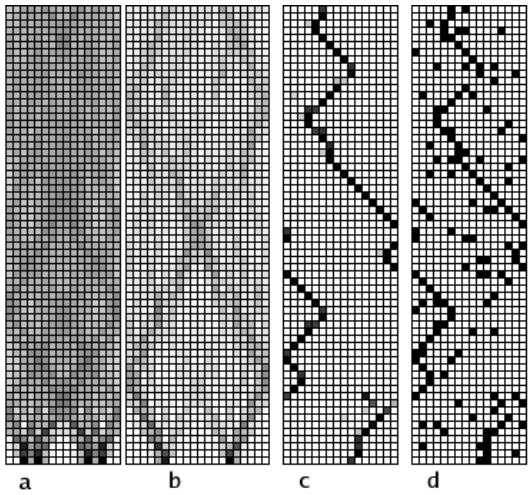


Figure 2. Plots of the quantum mechanical probabilities of quantum walkers. Plot (a) is a simulation for 4 walkers ($\theta = \pi/3$). Plot (b) shows two decohering walkers. The decohering effect is weak, with a choice of $X = 0.95$ ($\theta = \pi/3$). Plots (c) is a single walker simulation ($\theta = \pi/6$) with strong decoherence ($X = 0.3$), (d) is the "ancilla register" (the hits on field values) , for exactly the same run as (c). On all plots time runs upwards.

2. A Unitary Model on a Lattice: Quantum Walks

2.1. Unitary Model on a 1+1 Lattice

We analyse and simulate a dynamical (collapse) model [5,6,7] defined on a 1+1 lightcone lattice as seen in figure (1). The lattice is periodic in space (N vertices) and extend into the infinite future. The field variables $\alpha = 0, 1$ of the theory live on the links, so that on each link there is qubit Hilbert space $\mathcal{H}_\alpha$ with basis $\{|0\rangle, |1\rangle\}$. These field values, $\alpha = 0, 1$, can be thought of as occupation numbers, so that an "occupied" or "empty" state can be associated to each link and the diagonal links are the only worldline for propagation of particles. A quantum state $|\psi\rangle_\sigma$ on a spacelike surface σ is a vector in the 2^{2N} dimensional Hilbert space which is the tensor product of each $\mathcal{H}_\alpha$ of each link it cuts. A general state is a normalised function of the field variables $|\psi\rangle_\sigma = \sum_n^{n=2N} c_n |\alpha_1 ... \alpha_{2N}\rangle$.

In the standard lattice quantum field theory, there is a local evolution law at each vertex v_k embodied in a transfer R-matrix:

$$U = \begin{array}{c} \\ \nearrow \\ \\ \searrow \\ \searrow\nearrow \end{array} \begin{array}{c} \begin{array}{cccc} \nwarrow & & \nearrow & \nwarrow\nearrow \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & i\sin\theta & \cos\theta & 0 \\ 0 & \cos\theta & i\sin\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \tag{1}$$

The unitary evolution of the wave function from a spacelike surface σ to a surface σ', to the future of σ, is effected by applying all the R-matrices at the vertices between σ and σ' in an order respecting the causal order of the vertices.

2.2. Single and Multiple Quantum Walkers

Our model can be used to investigate Quantum Cellular Automata (QCA) in the spirit of Meyer [8]. A single particle QCA, with two-step translation invariance, is simulated by limiting the wave function to a sum over eigenstates with only one single $\alpha = 1$. We ran simulation of up to 20 walkers on lattice up to 10 vertex wide.

Let the walker position be the vertex k it is incoming to, so that the motion across a vertex changes its position to $k-1$ or $k+1$. The labelling of the vertices is shown in figure (1). The extra degree of freedom, which is whether the walker is on a left moving or right moving link, is embodied in the coin states $|01\rangle$ and $|10\rangle$, for each k, which we denote from now on as $\phi_{\nwarrow}$ and $\phi_{\nearrow}$. For a single walker, the unitary evolution of equation (1) is applied to the coin, and the movement of the walker across a vertex is conditioned by the result of the "coin tossing" operation:

$$\begin{pmatrix} \phi'_{\nearrow} \\ \phi'_{\nwarrow} \end{pmatrix} = \begin{pmatrix} i\sin\theta & \cos\theta \\ \cos\theta & i\sin\theta \end{pmatrix} \begin{pmatrix} \phi_{\nwarrow} \\ \phi_{\nearrow} \end{pmatrix} \tag{2}$$

The quantum walks of the previous section can be generalised to a system of interacting quantum walkers, by using the scattering matrix (1) on the basis states $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Similar 2-D lattice field theories have been identified as "quantum lattice gas automata" (QLGA) [8]. Four quantum walkers are shown in figure (1a) with $\theta = \pi/3$. Each link is represented by a cell whose darkness is positively proportional to the amplitude on the link.

3. A Causal Collapse Model on a Lattice and Decohering Quantum Walkers

Motivated by dynamical collapse models such as the CSL model, the lattice unitary model is made into a collapse model by introducing local "hits" driving the state into field eigenstates. The hits are partial spontaneous collapses on field eigenstates $\hat{\alpha}$ ("unsharp" projections), of the type:

$$J_{\alpha\hat{\alpha}} = \frac{\delta_{\alpha\hat{\alpha}} + (1 + \delta_{\alpha\hat{\alpha}})X}{\sqrt{1+X^2}} \quad 0 \leq X \leq 1 \text{ and } J(\alpha=0)^2 + J(\alpha=1)^2 = 1 \tag{3}$$

where α is the field variable on a particular link, and $\hat{\alpha}$ the centre of the hits. By analogy with the quantum mechanical rule, the hits are realised with probability, on each link:

$$\Pr(\hat{\alpha}) = ||J_{\hat{\alpha}}|\psi\rangle||^2 \tag{4}$$

where ψ is the wavefunction before the realisation of the hit around $\hat{\alpha}$. When both field values are realised on both links after each vertex v, the wavefunction is normalised.

In the quantum walk scenario, one can consider the hitting process as either the effect of partial measurement on the coin, or as the effect of an open system coupled to classical ancillae, with varying degrees of entanglement between the system and the ancillae. In the first scenario, depending on how strongly the walker's coin state is measured, we have a quantum walk or a near classical random walk, as shown in figure 2 (c). The classical limit is reached with strong, projective measurements. All quantum inteferences disappear and the record of the measurement outcomes singles out a particular classical path. In the second scenario, one can think of the walker coupled to local classical ancillae on each link, and the field configuration ("ancillae register") then labels the particular outcomes of measurements on the ancillae, as shown in figure 2 (d). The strength of the coupling to the ancillae determines whether the walk is unitary or classical. The quantum superpositions (interference of the walker(s)) introduced by the mixing effect of the unitary evolution (whose strength is determined by θ) are destroyed by the hits, whose sharpness is determined by X.

The simulations show clearly that the action of the J operators on each link can be thought of as either a fundamental new dynamics for the field, or as the decohering effect of a local coupling to the environment. Quantum walks exhibit clearly the trade off between the knowledge of which path the walker takes versus the interference effects.

Acknowledgements

I would like to thanks Fay Dowker for collaboration on this work.

References

[1] A. Bassi and G. C. Ghirardi, Dynamical reduction models, *Phys. Reports***379** (2003) 257-426.

[2] T. A. Brun, H. A. Carteret and A. Ambainis, Quantum random walks with decoherent coins, *Phys. Rev. A*, **67** (2003) 032304, [quant-ph/0210180].

[3] V. Kendon and B. Tregenna, Decoherence in discrete quantum walks, (2003) [quant-ph//0301182].

[4] V. Kendon and B. S. Tregenna, Decoherence can be useful in quantum walks, *Phys. Rev. A* **67** (2003) 042315 [quant-ph/0209005].

[5] F.Dowker and J.Henson, A spontaneous collapse model on a lattice,*J. Stat. Phys.***115** (2004) 1349, [http://xxx.lanl.gov/abs/quant-ph/0209051, quant-ph/0209051].

[6] F.Dowker and I.Herbauts, Simulating causal wave-function collapse models, *Class. Quant. Grav.* **21** (2004) 1–17, [http://xxx.lanl.gov/abs/quant-ph/0401075, quant-ph/0401075].

[7] F.Dowker and I.Herbauts, The status of the wave function in dynamical collapse models, *To be published in Found. of Phys. Lett.*, [quant-ph/0411050, http://xxx.lanl.gov/abs/quant-ph/0411050]

[8] David A. Meyer, From quantum cellular automata to quantum lattice gases, *J. Stat. Phys.* **85** (1996) 551-574, [http://xxx.lanl.gov/abs/quant-ph/9604003, quant-ph/9604003]

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Certain Aspects of Quantum Random Walk Asymptotics[1]

Ioannis Smyrnakis [a,2]
[a] *Technical University of Crete*

Abstract. In this work, the unconditional measurement quantum random walk (UMQRW) on the integers, Z, [3] is compared and contrasted with the more traditional model of quantum random walk (TQRW) [1,2] where the coin-walker system undergoes a unitary evolution and the walker position is read out by tracing the coin at the end of evolution. It is found that although both quantum walks display quadratic speedup in spread and distribution cutoffs in the asymptotic limit, in the case of the UMQRW there is a quantum version of the De-Moivre Laplace binomial approximation theorem, attributed to the use of many identical coins in the evolution process, and there is also a quantum version of the Chapman-Kolmogorov equation attributed to the possibility of interrupting the walk without spoiling its evolution.

Keywords. Quantum Computation,, Quantum Random Walks.

1. Introduction

A few years ago, the concept of a quantum random walk (qrw) was introduced [1]. In that model, the walker system was the position of a spin $1/2$ particle moving in the x-direction. The auxiliary coin system was the spin of this particle. The evolution of the system in unit time was determined by the unitary operator $U = \exp(-iS_z Pl/\hbar)$, which prescribes a displacement of the walker position by distance $\pm l$ conditional upon the spin. After repetition of this evolution procedure n time units, it was found that the walker position could deviate much faster (linear in n) than it could deviate in a classical stationary Markovian random walk ($\propto \sqrt{n}$).

Later on [2] and others studied a mathematical abstraction of this qrw on the integers Z. In their model, the walker Hilbert space was $H = l_2(Z) = span\{|n> \; | n \in Z\}$ and the two dimensional coin Hilbert space was $H_c = span\{|+>, |->\}$. The evolution operator $V : H_c \otimes H \rightarrow H_c \otimes H$ of the bipartite system consisted of a unitary coin reshuffling $U \otimes I$ which acts as a unitary operator U on H_c alone, followed by a conditional shift on the walker system $V_{cl} = (P_+ \otimes E_+ + P_- \otimes E_-)$. Here $P_\pm$ are projection operators and $E_\pm |n>= |n \pm 1>$. In this notation

[1]This work is supported by Pythagoras II.
[2]Correspondence to: Ioannis Smyrnakis, Department of Sciences, Division of Mathematics, Technical University of Crete GR-731 00 Chania, Crete, Greece. Tel.: +30 28970 37745 333; E-mail: smyrnaki@tem.uoc.gr.

$$V = V_{cl}(U \otimes I). \tag{1}$$

In their model, which we are going to call the traditional quantum random walk (TQRW) model, the walker density matrix after n evolutions steps is given by

$$\rho^{(n)} = \epsilon_{V^n}(\rho^{(0)}) = Tr_{H_c}(V^n \rho_c \otimes \rho^{(0)} V^{\dagger n}), \tag{2}$$

where $\rho^{(0)}$ is the initial walker density matrix and ρ_c is the initial coin density matrix. Here the number of steps n has a time of interaction interpretation. After the interaction time, the coin system is allowed to decouple from the walker system. This decoupling mathematically takes place by tracing out the coin degrees of freedom. The resulting walker density matrix is the n-steps evolved density matrix, and the diagonal elements of this density matrix give the position probability distribution of the walker after n evolution steps.

In the work of [2], among others, it was found that for large evolution time n, if the unitary coin reshuffling U is the Hadamard matrix and in the case the walker starts at the definite position 0, the asymptotic position distribution was confined in the interval $[-\frac{n}{\sqrt{2}}, \frac{n}{\sqrt{2}}]$.

There are two points that need to be stressed. The whole evolution process makes use of only one coin which is in sharp contrast to the classical random walk in which a coin toss is involved at every evolution step. Hence one can say that in the TQRW there is less randomness than in its classical counterpart. The second point is that it is not possible to interrupt the evolution of the TQRW. If the coin decoouples at some intermediate step then it is not possible to restart evolution.

A variation of the TQRW, that uses many identical coins and can be interrupted, has been introduced in [3]. Here the Hilbert spaces and the evolution operator are of the same form as in the TQRW, however the measurement procedure changes. An *unconditional* measurement (measurement without reading the result) is taken after a certain fixed number k of V applications, and the whole procedure is repeated n times, each time with a new identical coin. The parameter n measures the number of evolution steps and is equal to the number of coins needed to implement the walk. The unconditional measurement is realized through tracing in the coin system, thus decoupling the coin system after every k V applications.

In this unconditional measurement quantum random walk (UMQRW) model with parameter k, the walker density matrix after n evolution steps is given by

$$\rho^{(n)} = \epsilon^n_{V^k}(\rho^{(0)}), \tag{3}$$

where $\epsilon_{V^k}(\rho) = Tr_{H_c}(V^k \rho_c \otimes \rho V^{\dagger k})$. Hence for example $\rho^{(1)} = Tr_{H_c}(V^k \rho_c \otimes \rho^{(0)} V^{\dagger k})$, $\rho^{(2)} = Tr_{H_c}(V^k \rho_c \otimes \rho^{(1)} V^{\dagger k}) = Tr_{H_c}(V^k \rho_c \otimes \big[Tr_{H_c}(V^k \rho_c \otimes \rho^{(0)} V^{\dagger k})\big] V^{\dagger k})$, etc. In this type of evolution one uses *an identical coin at every evolution step*. Furthermore since the coin decouples at every evolution step, it is possible to interrupt the evolution so as to obtain a walker density matrix and then start it over again.

It is by now well known that the TQRW exhibits certain features that distinguish it from a classical random walk. First of all it spreads quadratically faster than a classical random walk. Furthermore the asymptotic distribution of the normalized position random variable exhibits cutoffs due to destructive amplitude interference beyond the cutoff [2], [5], [6]. As we are going to see, the UMQRW also exhibits these typically quantum

features, while at the same time classical random walk techniques can be used to analyze its behavior.

2. Quantum Characteristics of the Unconditional Measurement Quantum Random Walk

To analyze the evolution of the UMQRW, it is convenient to Fourier transform the walker position states

$$|\phi\rangle = \frac{1}{2\pi}\sum_{n\in Z} e^{-in\phi}\,|n\rangle\,, \tag{4}$$

since, if the coin state is $|c\rangle$,

$$V\,|c\rangle \otimes |\phi\rangle = (V(\phi)\,|c\rangle) \otimes |\phi\rangle \tag{5}$$

where $V(\phi) = e^{i\phi\sigma_3}U$ acts on the coin state $|c\rangle$ and σ_3 is the Pauli matrix.

In terms of these transformed states,

$$\epsilon_{V^k}(|\phi\rangle\,\langle\phi'|) = A(\phi,\phi')\,|\phi\rangle\,\langle\phi'|\,, \tag{6}$$

where $A(\phi,\phi') = Tr_{H_c}(V(\phi)^k\rho_c V(\phi')^{\dagger k})$ is the *characteristic function* of the UMQRW. Suppose now that the initial walker density matrix admits the following expansion in the $|\phi\rangle\,\langle\phi'|$ basis

$$\rho^{(0)} = \int_0^{2\pi}\int_0^{2\pi} d\phi d\phi'\,\rho(\phi,\phi')\,|\phi\rangle\,\langle\phi'| \tag{7}$$

Then if we use $L_c^{(n)}$ to denote the classical random variable that gives the position of the walker after n evolution steps, the probability the walker is in position m is given by

$$P(L_c^{(n)} = m) = \langle m|\,\rho^{(n)}\,|m\rangle = \frac{1}{(2\pi)^2}\int_0^{2\pi}\!\!\int_0^{2\pi} d\phi d\phi'\,\rho(\phi,\phi')A(\phi,\phi')^n e^{-im(\phi-\phi')} \tag{8}$$

where $\rho(\phi,\phi')$ is the Fourier transformed initial walker density matrix.

A direct computation of the moments of this distribution leads to the following result

$$\left\langle (L_c^{(n)})^s \right\rangle = \frac{1}{2\pi i^s}\int_0^{2\pi} d\phi\,\partial_\phi^s[\rho(\phi,\phi')A^n(\phi,\phi')]_{\phi'=\phi}. \tag{9}$$

This has as a direct consequence the following lemma

Lemma *The mean position of the quantum random walk after n evolution steps is of the form $\mu_n = K_1 n + \mu_0$, while the variance is of the form $\sigma_n^2 = K_2 n^2 + K_3 n + \sigma_0^2$, where K_1, K_2, K_3 depend on the coin density matrix, the initial walker density matrix and the tracing scheme, however they are independent of the number of evolution steps taken. μ_0 and σ_0^2 are the mean and variance of the initial position distribution.*

This means that if $K_2 \neq 0$ then $\sigma_n \propto n$ while for the classical random walk $\sigma_n \propto \sqrt{n}$. However there are in general sharp resonances in which $K_2 = 0$ and hence the UMQRW has classical spread. [4]

Since

$$\left\langle (L_c^{(n)})^s \right\rangle = \frac{n^s}{2\pi} \int_0^{2\pi} d\phi \rho(\phi, \phi')[-i\partial_\phi A(\phi, \phi')]^s_{\phi'=\phi} + O(n^{s-1}). \tag{10}$$

the leading order contribution to the moments with respect to the number of steps depends only on the *asymptotic characteristic function*

$$h(\phi) = -i\partial_\phi A(\phi, \phi') = ImTr_{H_c}(V^{k\dagger}(\phi) \left(V^k(\phi)\right)' \rho_c). \tag{11}$$

This function determines completely the asymptotic distribution of the normalized random variable $L_c^{(n)}/n$. In fact the limiting form of the moments of $L_c^{(n)}/n$ gives us the following distribution theorem.

Theorem *The sequence of random variables* $\frac{L_c^{(n)}}{n}$ *converges weakly to the random variable* $Y = h(\phi)$, *where* ϕ *is now an angular random variable with values on the interval* $[0, 2\pi]$ *and measure* $\frac{1}{2\pi}\rho(\phi, \phi)$.

This is similar in form to Theorem 1 in [6]. Nevertheless there are also significant differences since the measure in our case depends on the initial walker density matrix and not on the coin density matrix, and furthermore we do not need any extension of the sample space by a discrete set as was required in [6]. Also it should be mentioned that the first to obtain such asymptotic distribution theorems for the TQRW was [5].

If $h(\phi)$ happens to be a constant then we have classical spread.

As an example, suppose that the walker is initially at position 0, the parameter $k = 2$ and consider the particular case where the coin is initially in the state $|+>$ while the coin reshuffling unitary matrix U is of the form $U = e^{i\frac{\pi}{4}\sigma_2}$. Then the asymptotic characteristic function is $h(\phi) = -cos2\phi$ and $\rho(\phi, \phi) = 1$. A direct application of the above theorem gives us that the asymptotic distribution of the random variable Y is given by

$$P(y_1 \leq Y \leq y_2) = \frac{1}{2\pi} \int_{y_1 \leq h(\phi) \leq y_2} d\phi = \frac{1}{\pi} \int_{y_1}^{y_2} \frac{1}{\sqrt{1-y^2}} dy. \tag{12}$$

This is an asymptotic distribution of the double horn shape (an arcsin distribution) with cutoffs at $Y = \pm 1$ while the support of the distribution was expected to lie in the interval $[-2, 2]$ since $k = 2$. This distribution appeared also in [7], although in a different context.

3. Differences Between the Unconditional Measurement Quantum Random Walk and the Traditional Quantum Random Walk

One of the characteristics of the UMQRW that is not true for the TQRW is that it utilizes many identical coins. This leads one to expect stronger universality theorems because of the enhanced similarity to the classical random walk which also utilizes many coin tosses.

To see this, suppose one starts with a walker at position 0 and applies ϵ_{V^2} evolution. For a general coin density matrix and a general U, the asymptotic distribution function is of the form $h(\phi) = \mu - \sqrt{2}\sigma cos(2(\phi + \alpha))$ where the parameters μ and σ are identified with the mean and the standard deviation of the asymptotic random variable $Y = h(\phi)$.

Since ϕ is a uniform random variable on $[0, 2\pi]$ the random variable Y satisfies the following theorem which should be thought of as the ϵ_{V^2} quantum De Moivre-Laplace theorem:

Theorem *Whatever the initial coin density matrix and whatever the unitary coin reshuffling matrix for the ϵ_{V^2} quantum random walk, if the asymptotic characteristic function $h(\phi)$ is not constant then the normalized random variable $Y_N = \frac{Y-\mu}{\sqrt{2}\sigma}$ is distributed according to the distribution $P(y_1 \leq Y_N \leq y_2) = \frac{1}{\pi}\int_{y_1}^{y_2} \frac{1}{\sqrt{1-y^2}} dy$.*

A second characteristic of the UMQRW not present in the TQRW is the fact that one can stop the walk and then restart it again. This leads to a compatibility equation which is a form of the Chapman-Kolmogorov equation. Suppose that by $\rho^{(s_0,n)}$ we denote the ϵ_{V^k} walker density matrix after n evolution steps if the walker starts with density matrix $\rho^{s_0} = |0\rangle\langle 0|$ and by $\rho^{(s,n)}$ the corresponding matrix if the walker starts at a general density matrix ρ^s. The quantum version of the Chapman-Kolmogorov equation is

$$\rho^{(s,n_1+n_2)}_{m_1,m_2} = \sum_{\sigma_1,\sigma_2 \in Z} \rho^{(s_0,n_2)}_{m_1-\sigma_1,m_2-\sigma_2} \rho^{(s,n_1)}_{\sigma_1,\sigma_2}. \tag{13}$$

If we take $n_2 = 1$ then this equation takes the form of a recurrence relation that gives $\rho^{(s,n+1)}$ in terms of $\rho^{(s,n)}$. The transition density matrix $\rho^{(s_0,1)}$ does not depend on the initial density matrix ρ^s of the walker system, so it is similar to a set of classical transition probabilities in a classical stationary Markovian random walk. This equation can be used to derive differential equations for the asymptotic walker distribution by rewriting it in an appropriate semi-continuous limit, in a way somewhat similar to the derivation of the diffusion equation from a classical random walk recurrence relation.

Suppose for example that $\rho_c = |c><c|$ where $|c> = \cos\theta|+> + i\sin\theta|->$, that $\rho^s = \rho^{s_0} = |0><0|$, that $U = e^{i\frac{\pi}{4}\sigma_2}$ and that $k = 2$. If we keep $m_2 - m_1 = m'$ discrete while making m_1, n continuous through the introduction of scales Δy, Δt respectively with $\Delta y/\Delta t \to D$ then we obtain the following set of partial differential equations

$$\frac{1}{D}\frac{\partial}{\partial t}\rho_{m'}(y,t) \tag{14}$$

$$= \frac{1}{2}\left[e^{2i\theta}\frac{\partial}{\partial y}\rho_{m'+2}(y,t) + e^{-2i\theta}\frac{\partial}{\partial y}\rho_{m'-2}(y,t)\right] \tag{15}$$

subject to the boundary conditions $\rho_{m'}(y,0) = \delta_{m',0}\delta(y)$. This admits the solution

$$\rho_{2m'}(y,t) = \frac{1}{\pi}e^{-2im'\theta}T_{m'}(-\frac{y}{Dt})\frac{1}{\sqrt{1-\frac{y^2}{D^2t^2}}}, \tag{16}$$

$$\rho_{2m'+1}(y,t) = 0 \tag{17}$$

where $T_m(y)$ are the Chebychev polynomials. Note that in the particular case $m' = 0$ we get $\rho_0(y,t) = \frac{1}{\pi}\frac{1}{\sqrt{1-\frac{y^2}{D^2t^2}}}$ and this distribution corresponds to the diagonal elements of the density matrix. This is the continuous time version of the asymptotic ϵ_{V^2} quantum random walk position distribution (12). It bears the same relation to the distribution (12), as the Wiener process has to the position distribution of an n step discrete stationary Markovian classical random walk.

There is however one difference that is important between the classical diffusion equation and the set of differential equations (14). Here we have kept $m_2 - m_1 = m'$ *discrete*. This is necessary to be able to take a continuum limit. One way to understand this necessity is by the fact that along the diagonal of the density matrix we have a probability density function that integrates to 1. If the above difference was continuous, then this would integrate to zero, since the diagonal has no width, hence its area is zero. This renders necessary the semi-continuous limit.

4. Conclusion

It has now become apparent that an alternative version of the TQRW, the UMQRW, behaves also in a quantum manner, while at the same time it bears a closer relation to the classical random walk. It behaves in a quantum manner because generically, for $k > 1$, it spreads quadratically faster than a classical random walk, and because asymptotically it exhibits the typical quantum cutoffs before maximal possible spread. It is closer to a classical random walk because it utilizes many identical coins, one for each evolution step. In fact the $k = 1$ UMQRW model gives precisely classical random walks. Furthermore there is a quantum version of the Chapman-Kolmogorov equation satisfied by the evolved walker density matrix. This leads to a set of partial differential equations in the semi-continuous limit for the diagonals of the density matrix.

References

[1] Y. Aharonov, L. Davidovich and N. Zagury. *Quantum Random Walks.* Physical Review A **48** 1687 (1993).

[2] A. Ambainis, E. Bach, A. Nayak, A. Vishwanath and J. Watrous, *One dimensional quantum walks*, Proceedings of the 33th Annual ACM Symposium on Theory of Computing, 37-49, (2001)

[3] Anthony J. Bracken, Demosthenes Ellinas and Ioannis Tsohantjis, *Pseudo Memory Effects, Majorization and Entropy in Quantum Random Walks*, J. Phys. A : Math. Gen. 37 (2004) L91-L97.

[4] Demosthenes Ellinas and Ioannis Smyrnakis, *Asymptotics of Quantum Random Walk Driven by Optical Cavity,* To appear in J. Opt. B: Quantum Semiclass.

[5] Norio Konno, A New Type of Limit Theorems for the One-Dimensional Quantum Random Walk, quant-ph/0206103

[6] Geoffrey Grimmett, Svante Janson and Petra Scudo, Weak limits for quantum random walks, Phys. Rev. E, 69, 026119 (2004)

[7] Norio Konno, Continuous-Time Quantum Walk on the Line, quant-ph/0408140.

Acknowledgements

I would like to thank Prof. D. Ellinas for useful discussions.

This page intentionally left blank

Chapter 4

Implementations of Quantum Information Processing

This page intentionally left blank

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

From Entanglement to Quantum Key Distribution

Hannes Hübel [a,1] and Anton Zeilinger [a,b]
[a] *Institut für Experimentalphysik, Universität Wien*
[b] *Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences*

Abstract. This paper concentrates on the experimental applications of entangled photons in the field of quantum information. The production of entangled photons is reviewed, focusing on polarization entanglement. Transmission of photons takes place either via optical fibers or free space. Those channels are compared, together with different detection methods for single photons and entanglement. The value of entangled photons is highlighted by discussing Bell experiments and quantum teleportation. Furthermore the first experimental realization of a quantum computer based on photonic cluster states is presented. Free space distribution of entanglement is introduced as a possible way for long distance quantum communication. Various experimental schemes for quantum cryptography are discussed. Weak coherent pulses and entanglement based implementations are compared. A full QKD experiment using polarization entanglement is reviewed in detail.

Keywords. Entanglement, photons, QKD, quantum information processing, quantum computing

1. Introduction

Quantum physics has excited and fascinated people for the last century. However it is only in the last decades that novel application were found, that would make explicitly use of the fundamental principles of quantum mechanics. The field of quantum information emerged.

In quantum information, single quantum states, called qubits, are the carrier of information. This allows extending the rules of quantum physics to information science. In contrast to classical bits, the quantum mechanical approach can exhibit features like superposition or entanglement of qubits and therefore of the information, which gives rise to many interesting and valuable applications. Today many experimental implementations exist in the laboratory and some are already on the way to commercialization.

This paper is not meant to be a thorough review of the field, but rather a tour from the production of quantum entanglement to its very valuable and imminent application of quantum cryptography, focusing on the importance of entanglement in quantum information experiments. Along the road some of the now "classic" experiments are en-

[1]Correspondence to: Hannes Hübel, Institut für Experimentalphysik, Universität Wien, Boltzmanngasse 5, 1090 Wien, Austria. Tel.: +43-1-4277 51227; Fax: +43-1-4277 51229; E-mail: hannes.huebel@univie.ac.at.

countered together with tools needed to manipulate entangled photons. It also includes some recent results to give a flavor of in which direction the field is heading. The way is roughly divided into three sections.

In the first part we give an overview on the production of entangled photons. Different setups are presented, particularly focusing on polarization entanglement. After some description of the possible quantum channels used for transmission, the detection of single photons is reviewed. Finally some experimental methods to verify entanglement are introduced.

The second part deals with some important experimental breakthroughs in quantum information achieved over the last years using entangled photons. It includes some new experiments which have proved the possibility of long distance quantum communication applications together with the experimental realization of a novel technique in quantum computation.

Quantum cryptography is at last presented. It is the first quantum information application which is nearly ready to be commercially exploited. The basic ideas are presented together with their experimental implementations. An experiment using polarization entangled photons for quantum cryptography is discussed in detail.

2. Entangled Photons

2.1. Production of Entangled Photons

Spontaneous parametric downconversion (SPDC) is nowadays to most used method for producing entangled photons. It relies on the nonlinearity of the polarization in some crystal structures. An expansion of the vectorial components of the polarization to low orders gives:

$$P_i = \chi_{ij}^{(1)} E_j + \chi_{ijk}^{(2)} E_j E_k + \chi_{ijkl}^{(3)} E_j E_k E_l + \cdots \tag{1}$$

where $\chi^{(2)}$ is the first nonlinear factor.

SPDC is a special case of a $\chi^{(2)}$ nonlinear process in which a pump photon decays into two photons called signal and idler photon. The interaction must conserve energy and momentum:

$$\begin{aligned} \omega_{pump} &= \omega_{signal} + \omega_{idler} \\ \mathbf{k}_{pump} &= \mathbf{k}_{signal} + \mathbf{k}_{idler} \end{aligned} \tag{2}$$

These relations are also called phase matching conditions. The overall efficiency of the SPDC process is strongly dependent on a perfect matching. But even with perfect phase matching the conversion efficiency of a pump photon into a pair of low-energy down-converted photons is still very small, in the order of 10^{-9}. Therefore, optical pumping needs to be sufficient high in order to generate down-conversion photons at an observable rate. One further distinguishes the down conversion into two schemes depending on the polarization of the signal and idler photon:

- Type-I: signal and idler photon have the same polarization.
- Type-II: signal and idler photon have orthogonal polarization.

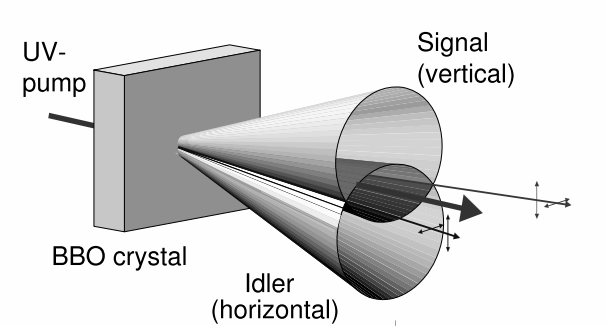


Figure 1. Production of polarization entanglement. Orthogonally polarized photons are emitted along cones from a BBO crystal. The photon pair from the intersection points is entangled.

The strong momentum and time correlations of the downconverted photons can be used to entangle them.

Due to the momentum conservation, the signal and idler have opposite momentum components perpendicular to the pump beam. Spatial filtering can be used to select only two possible pairs of momentum components (paths). Using beamsplitters, the paths are combined and any knowledge which photon took which path is erased [2]. The first experimental test of this kind of entanglement was performed by Rarity and Tapster [3].

Another scheme is the Franson-type energy-time-entanglement [4,5]. It is based on the fact that the down-converted photons are created simultaneously. When both photons travel through unbalanced but identical interferometers, coincidence events are only recorded when both photons take the same path (long-long or short-short) through the interferometers. By making the long-long and short-short possibilities indistinguishable one will again obtain entanglement.

2.1.1. Polarization Entanglement

The photons can also be entangled in other degrees of freedom, such as polarization entanglement [1]. In general the phase matching condition is only fulfilled for a certain direction and polarization at a given frequency. In practice, for type-II SPDC, photons at a given wavelength are emitted along cones with different polarizations on each cone, as seen in figure 1.

By adjusting the angle between the pump direction and the optical axis of the crystal the two cones can be brought to intersect. At the intersection points one cannot tell -even in principle- from which cones the photons originated and is thus left with an unpolarised state of light. But one does know that the photons at the two intersection points must have opposite polarization (type-II). This leads to a polarization entangled state of the form:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|V\rangle_1|H\rangle_2 + e^{i\theta}|H\rangle_1|V\rangle_2) \tag{3}$$

where H and V denote the horizontal and vertical polarization of the photons.

Any distinguishability would lead to a reduction of the entanglement. Therefore, care must be taken to include only photons which are indistinguishable:

1. Signal and idler must be frequency degenerate. A difference in photon energy would distinguish them. Narrow spectral filters have to be applied for high-quality entanglement due to the natural broad bandwidth of the SPDC generated photons.
2. Photons must only be selected from the intersection points, otherwise their position would betray their polarization state. Careful spatial filtering is necessary.
3. The crystal itself generates spatial and temporal walk-off effects due to the fact that the signal and idler photons have different group velocities in the birefringent medium. The walk-off must be compensated with additional optical elements.

The phase of the entangled state can be changed by varying the optical pathlength for one of the photons. Together with polarization rotation of one photon, any of the four maximally entangled Bell states can be produced:

$$
\begin{aligned}
|\psi^-\rangle &= \frac{1}{\sqrt{2}}(|V\rangle_1|H\rangle_2 - |H\rangle_1|V\rangle_2) \\
|\psi^+\rangle &= \frac{1}{\sqrt{2}}(|V\rangle_1|H\rangle_2 + |H\rangle_1|V\rangle_2) \\
|\phi^-\rangle &= \frac{1}{\sqrt{2}}(|V\rangle_1|V\rangle_2 - |H\rangle_1|H\rangle_2) \\
|\phi^+\rangle &= \frac{1}{\sqrt{2}}(|V\rangle_1|V\rangle_2 + |H\rangle_1|H\rangle_2)
\end{aligned} \tag{4}
$$

2.1.2. Asymmetric Setups

There is a large variety of setups designed to produce polarization entanglement [6,7,8]. Here we mention the class of asymmetric setups which are designed to provide photons at commercially used telecom wavelengths, ideally suited for long distance transmission. For low attenuation in optical fibers, the wavelength of photons should be in the infrared around 1550 nm. However at those wavelengths the efficiency of single photon detectors is low (10-15%) compared to detectors operating in the visible regime (40%-60%). To take advantage of low attenuation and high detection, asymmetric sources have been proposed [9], where the signal photon is produced in the visible or near infrared, while the idler photon is produced at telecom wavelengths while obeying the phasematching conditions in eq. 2. In order to keep the attenuation for the signal small, it should be detected near the source. To obtain polarization entanglement with non-degenerate photons the geometry of the setups has to be modified [10,11].

2.1.3. Quasi-Phasematching

In the case of asymmetric sources it is very hard to find a material with just the right dispersion properties to satisfy the phasematching conditions due to the large difference in the wavelengths. A new technique, the so called quasi-phasematching, allows the creation of photons at nearly any wavelength.

In this technique the phase-mismatch, which would lead to very low conversion efficiencies, is compensated for by an artificial wave-vector. This is achieved by modulating the sign of the nonlinearity within the crystal along the propagation direction of the pump light. To produce such a periodically poled crystal, a high voltage with alternating

polarity is applied along one side of the crystal. Underneath each electrode a single crystal domain will be formed pointing in the opposite direction with respect to its nearest neighbors. The phase matching conditions are altered in the following way:

$$\mathbf{k}_p = \mathbf{k}_s + \mathbf{k}_i + \frac{2\pi}{\Lambda} \tag{5}$$

where Λ is the poling period. One can see that by being able to freely choose Λ, the range of interacting wavelengths is greatly enhanced.

Furthermore, by careful temperature control of the crystal, the emission can be limited to a direction close to the pump beam. In this co-linear geometry the signal and idler, emitted along the pump beam, show very high collection efficiencies for single mode fibers.

2.2. Transmission of Entangled Photons

2.2.1. Optical Fiber

The transmission of photons over optical fibers has the advantage that the light is guided in the fiber and can therefore be directed to any place a fiber can reach. The existing fiber infrastructure can be used, with exception of classical repeater stations which would destroy the quantum state (an arbitrary quantum state cannot be copied). The disadvantages lie mainly in the absorption and decoherence effects:

- Up to now, only photons around the telecom wavelength of 1550 nm have low attenuation ($\sim$0.2 dB/km)
- Chromatic dispersion causes a broadening of the pulse, which poses problems when time entanglement is used.

The following effects cause a reduction of entanglement of a polarization entangled state:

- The fiber introduces an arbitrary rotation depending on environmental parameters. If the changes are slow, an active compensation can be used.
- Polarization mode dispersion, caused by the coupling between orthogonal polarization modes, leads to a loss of coherence.

2.2.2. Free-Space

Free-space distribution offers advantages when using polarization entanglement, since none of the decoherence mechanisms are present in air. Furthermore, no additional channels are needed and the transmission can be between any points with a free line of sight between them. To minimize collection losses, the apertures of the receiving telescope should be as large as possible. Active tracking of the beam is also advantegeous. Further restrictions are the operation with low background light, preferably at night, although day-time operation can be achieved with additional filtering [12] and the dependence on meteorological conditions.

2.3. Detection of Entangled Photons

2.3.1. Single Photon Detectors

Since the information is actually carried by single particles, one needs to be able to address and measure those carriers individually. To detect single photons, avalanche photo diodes (APD) are often used. APDs are driven above breakdown voltage (Geiger mode). If a photon strikes the semiconductor material, an avalanche of carriers is triggered and the subsequent current can be measured. The voltage is then reduced as fast as possible below breakdown to prevent damage, the APD is then ready again to detect the next photon (after a few ns dead-time). Silicon APDs are nowadays widely used for detection below 1μm. They feature high efficiency (40-60%), low dark counts ($\sim$100Hz) and a large count rate (few MHz). For photons at longer wavelengths, especially at 1550 nm, the choice falls on Indium Gallium Arsenide APDs with efficiencies around 15%. Unfortunately those detectors have high dark count rates. To keep the dark counts at an acceptable level, the APD is only charged above breakdown when a photon is expected to arrive at the detector (gated mode). This requires an synchronized external trigger signal.

A novel approach could remedy the bad detection characteristics at 1550 nm. During the process of up conversion a single photon at low wavelengths interacts with a strong laser field in a crystal to produce a photon with a higher energy. First studies of upconversion show an overall efficiency of 40% [13,14].

Other detector schemes at 1550 nm are based on bolometric devices where the photon energy is dissipated into heat. This technique is sensitive enough to allow for photon number discrimination which is not possible with APD's. Their drawback is however the operation at cryogenic temperatures (few K) and the resulting technical complications.

2.3.2. Verification of Entanglement

Three different methods to quantify entanglement in two-particle systems are mainly used in experiments:

1. Measurement of the Bell parameter. A violation of the Bell inequality implies the existence of an entangled state (more details are given in the next section).
2. Visibility measurements are a quick way to determine the entanglement. The coincidence rate as a function of polarization of a photon (the polarization of second photon being fixed) shows sinusoidal fringes. Measured in two conjugate basis, the visibility of these fringes should be unity for an entangled state.
3. The most complete way to show if a state is entangled is by using state tomography. Various polarization measurements are combined to give a full representation of the density matrix of the state (ρ_{exp}). The degree of entanglement is often described by the parameter F. The fidelity F measures the overlap of the experimentally determined density matrix with the desired entangled state: $F = \langle\Psi|\rho_{exp}|\Psi\rangle$. Although this method gives the best results, the experimental complexity increases rapidly with the number of particles in the system.

2.4. Bell Theorem

In the famous EPR paper [16], Einstein questioned the completeness of the quantum mechanical description of physical reality on the example of an entangled quantum system.

In Bohm's version [15] of the EPR paradox a spin 0 particle decays into two spin 1/2 particles. The particles have opposite spins, but no particle carries a defined spin. The total state is entangled and can be written as:

$$\frac{1}{\sqrt{2}}(|\uparrow\rangle_1|\downarrow\rangle_2 + |\downarrow\rangle_1|\uparrow\rangle_2) \tag{6}$$

Measurement of the spin of one particle immediately implies knowledge about the spin of the other particle. Since the two particles can be separated over arbitrary distances, the EPR paper argues that the measurement of particle 1 could not have influenced particle 2.

EPR concluded that the quantum mechanical description of reality was not complete. They hinted that an additional formalism was needed. One of the most obvious extension would be a local hidden variable theory.

In 1964, John Bell [17] showed that for measurements of correlated quantities the predictions of quantum mechanics differ from a local hidden variable theory. His famous Bell inequality puts a bound on the correlations of a local hidden variable theory, but is violated by quantum mechanics.

To include experimental inefficiencies, the CHSH inequality [18], a modified Bell inequality, is used:

$$S_{CHSH}(\alpha, \alpha', \beta, \beta') \equiv |E(\alpha, \beta) - E(\alpha, \beta')| + |E(\alpha', \beta') + E(\alpha', \beta)| \leq 2 \tag{7}$$

where $|E(\alpha, \beta)|$ is the expectation value of simultaneously measuring the polarization of one particle in the direction α and the other particle along β.

The expectation values are calculated from the coincidence clicks between the two separate detectors. The number of coincidences is denoted by $C_{ij}(\alpha, \beta)$ where i and j stand for two possible outcomes of a polarization measurement. The coincidence count are then combined to yield $E(\alpha, \beta) = [C_{++}(\alpha, \beta) + C_{--}(\alpha, \beta) - C_{+-}(\alpha, \beta) - C_{-+}(\alpha, \beta)]/N$, where N is the sum of all coincidences.

Calculating the Bell parameter S for a quantum mechanical entangled system at the angles (0°, 22.5°, 45°, 67.5°) yields:

$$S_{CHSH} = 2\sqrt{2} \tag{8}$$

which is clearly violating the predictions of a local hidden variable theory.

However several experimental loopholes exist which could still allow a local hidden variable theory to enter by the back door.

- Locality loophole: the measurements on particle 1 and 2 have to be space-like separated so that no information (at the speed of light) could have been exchanged.
- Detection loophole: during the experiment an assumption has to be made that the subset of detected particles represents a fair sample of the total particles detected. Strictly speaking, the detected particles could, however unlikely, conspire against the experimenter and show a violation of the inequality, whereas if the whole set could be measured no violation would occur.
- Random settings (or free will) loophole: the choice of measurement settings has to be completely random. Otherwise the correlations in the settings would be predictable prior to measurement.

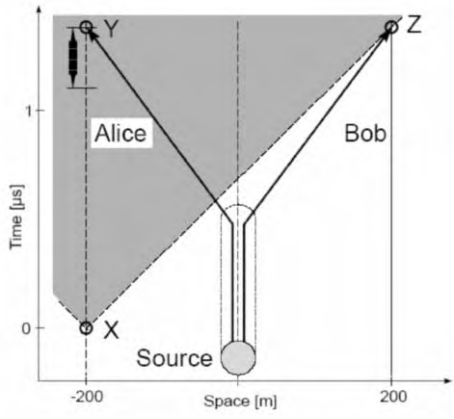


Figure 2. Space-time diagram of the Bell experiment, from [20]. The shaded area is space like separated from the detection event at Bob. Alice's choice of angle and detection are well within the shaded zone. The black bar at Alice indicates the duration of the measurement process.

3. Fundamental Experiments with Entangled Photons

Having dealt with the basics of the production of entangled photons and their detection, we now turn to the various uses. These range from fundamental tests of the principles of quantum physics to powerful applications in quantum information.

3.1. Bell Experiments

Experimental verification of the Bell's inequality was started by Freedman and Clauser in 1972 [21]. The first experiment to close the locality loophole was by Aspect in 1982 [22]. In his setup the two analyzers were separated by 12 m and switching was achieved in less than 10 ns providing a space-like separation of the switching events. A violation of the CHSH inequality was found. Although the analyzers were switched faster than signals could travel between them, the switching was performed periodically. One could therefore still argue that the future positions of the analyzer could have been predicted and known prior to the measurement.

The experiment which eventually closed both the locality and the randomness loophole at the same time was performed in 1998 [20]. Polarization entangled photons, produced via SPDC, were sent to two analyzer stations separated by 400 m. The settings on the analyzer to test the Bell inequality were chosen by a physical random number generator. The number generator was fast enough to supply the settings only 100 ns before the detection. Since the separation between the two analyzers was 1.3 μs (at the speed of light) the settings could not influence each other in any way. Their space like separation is depicted in figure 2. Due to the true randomness of the number generator no correlation in the analyzer settings was present. A Bell parameter of $S = 2.73 \pm 0.02$ was measured, showing a striking violating of the CHSH inequality by 37 standard deviations.

Due to the low efficiency of the single photon detectors the closure of the detection loophole could not yet be realized with photons. It was closed in an experiment with trapped ions [19] which have detection efficiencies close to 100%. Due to the proximity of the ions however, the locality loophole remained.

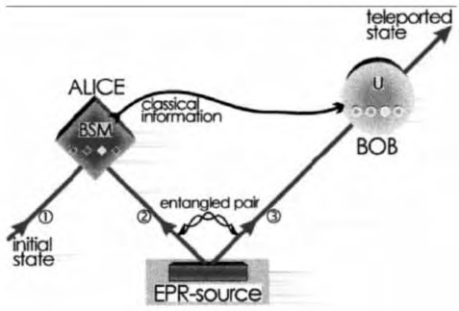


Figure 3. Schematic of the teleportation. A Bell measurement is performed on Photon 1 and an ancilla (photon 2) which is part of an entangled pair. After the measurement the original quantum state of photon 1 is carried by photon 3 (up to some unitary transformation).

One of the remaining challenges in the foundation of quantum mechanics is to perform a Bell experiment which closes all the loopholes at once.

3.2. Teleportation

Let us consider the following problem: Alice wants to transfer an unknown quantum state (qubit) $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ to Bob. Let us further assume that she cannot directly send the qubit. Any measurement on the qubit would lead to only a partial knowledge of the state and to the projection of the qubit into the measurement basis. How can she reliably tell Bob her quantum state? The teleportation scheme [33] allows her to do that.

3.2.1. Teleportation Protocol

For the scheme to work, a pair of entangled ancilla qubits (2, 3) are needed which are shared by Alice and Bob, respectively, as seen in figure 3. The pair is in the following state:

$$|\psi\rangle_{23} = \frac{1}{\sqrt{2}}(|0\rangle_2|1\rangle_3 - |1\rangle_2|0\rangle_3) \tag{9}$$

The total state of the system can be described by a tensor product of the subsystems:

$$|\psi\rangle_{123} = |\psi\rangle_1 \otimes |\psi\rangle_{23} \tag{10}$$

One can now re-express the joint state of qubit 1 and 2 using the four maximally entangled Bell states ($|\psi^+\rangle$, $|\psi^-\rangle$, $|\phi^+\rangle$ and $|\phi^-\rangle$), since they form a complete basis in the two-qubit Hilbert space. The total state can be rewritten as (normalization factors omitted):

$$\begin{aligned} &|\psi\rangle_{123} = \\ &|\psi^-\rangle_{12}(-\alpha|0\rangle_3 - \beta|1\rangle_3) + |\psi^+\rangle_{12}(-\alpha|0\rangle_3 + \beta|1\rangle_3) \\ &+|\phi^+\rangle_{12}(\alpha|1\rangle_3 - \beta|0\rangle_3) + |\phi^-\rangle_{12}(\alpha|1\rangle_3 + \beta|0\rangle_3) \end{aligned} \tag{11}$$

Alice can project her qubits in the Bell basis by performing a so called Bell state measurement. Once she has an outcome, the qubit at Bob's side will be in the state

according to eq. 11. For example, if Alice measures $|\psi^-\rangle$ then Bob's qubit is the exact copy of $|\psi\rangle_1$ (the overall phase is of no consequence). Similarly if Alice measures any of the other Bell states, Bob only needs to perform some unitary operation (phase change or flip of the basis) to recover $|\psi\rangle_1$ on his qubit.

Here is one important point: Since the outcome of the Bell-state measurement is completely random, Bob needs some classical signal from Alice telling him her measurement outcome. The teleportation cannot be used to send superluminal information.

Furthermore, Alice never learns the state of her system during the procedure and after the Bell measurement, her qubit will be destroyed. There is no duplication of the quantum state in compliance with the no-cloning theorem [62].

3.2.2. Bell State Analyzer

The difficulty lies in performing the Bell measurement 100% efficient. It was shown that in the case of photons, this is not possible using linear optics alone [32]. Up to now, using linear optics, only probabilistic schemes with a maximum efficiency of 50% exist.

It is however, straightforward to build a Bell analyzer which works only in 50% of the time (i.e. one can distinguish 2 out of the 4 Bell states). It is based on two-photon interference at a beamsplitter. In an early experiment in 1987 [31], two indistinguishable photons were combined at a beamsplitter to observe an effect based on two-photon interference. The interferenc would cause the photons always to exit through the same output port. The observed reduction in the coincidence rate is also called the Houng-Ou-Mandl dip.

To discriminate bell states via two-photon interference, one has to make additional use of the symmetry properties of the wavefunction. Photons are bosons and hence have an overall symmetric wavefunction (spin and spatial part combined). If the spatial part of the wavefunction is symmetric, both photons leave through the same port. Only an antisymmetric spin contribution ($|\psi^-\rangle$) will lead to a detection of the photons at different outputs. If two photons are measured in coincidence behind a beamsplitter then the pair was in the $|\psi^-\rangle$ state.

Looking at the remaining Bell states, one sees that the $|\psi^+\rangle$ contains photons with orthogonal polarizations, which should be easy to distinguish. By placing two additional polarizing beamsplitters (PBS) behind each output port, the photons will go different ways at the PBS if they are indeed in the state $|\psi^+\rangle$.

In this way a more complete Bell state analyzer can be realized.

3.2.3. First Experimental Realization

The first teleportation using photons was achieved in 1997 [30], the set-up is shown in figure 4.

A pulsed laser creates two pairs of entangled photons $|\psi^-\rangle_{14}$ and $|\psi^-\rangle_{23}$, the latter being the ancilla needed for the teleportation. The entanglement between 1 and 4 is not important as photon 4 is only used as a trigger. Photon 1 passes a polarizer to prepare it in the 'to be teleported' state. Alice then performs a Bell measurement with photon 1 and 2. In this setup only the singlet $|\psi^-\rangle_{12}$ state could be identified (when detectors f1 and f2 fired together) and hence the teleportation had a 1/4 chance of success. Photon 3, transmitted to Bob, carries the teleported state.

The delay of the second pump pulse could be varied to obtain maximal overlap at the beamsplitter. Teleportation is only possible in a certain region of the delay when both

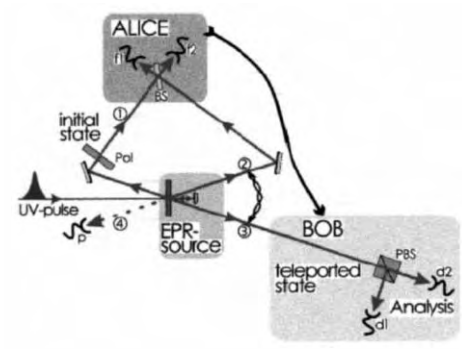


Figure 4. Setup of the first teleportation (from [30]): a SPDC source is used to produce two photon pairs. Alice prepares photon 1 in a specific state. The Bell measurement is performed at the beamsplitter on photon 1 and 2. After measurement, photon 3 carries the prepared state of photon 1 and is analyzed by Bob.

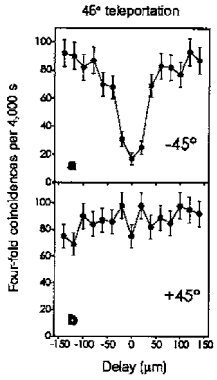


Figure 5. Results of teleporting a 45° state. Measuring in the -45° orientation shows a dip in the teleportation region. The rate for the +45° orientation remains constant, as expected.

photons arrive at the same time. This is however not enough to guarantee indistinguishability since the wave packets of photon 1 and 2 are much smaller (50 fs) than the pump pulse (200 fs). A coincidence detection with time resolution better than 50 fs could therefore in principle identify the photon pairs. To avoid this, spectral filters (4 nm) were used to give the photons a much longer coherence time (500 fs).

In order to show that teleportation has indeed taken place, photon 1 was prepared in state +45° and photon 3 was analyzed in the 45° basis. Events were only taken into account when all four (f1, f2, p and d) detectors fired. The delay of the second pump pulse was swept through the teleportation region and it was expected to see a dip at the detector measuring photon 3 in -45° . At the same time one would expect a rise of coincidence counts in the +45° basis. However due to the fact that teleportation is only working in 25% of the time the probability to obtain a four-fold coincidence with the +45 ° detector stays constant independently of the delay being set within the teleportation region or not.

Figure 5 shows the obtained data and confirms the teleportation of a +45 ° state. To show that the experiment indeed works for an arbitrary quantum state, teleportation was also performed with H, V and circular polarized photons. The fidelity of the teleported photons was found to be 70%.

3.3. Long Distance Teleportation

For future applications such as the quantum repeater or quantum communication methods in general, it is important to show the feasibility of long distance quantum teleportation. A high-fidelity teleportation over 600 m was demonstrated in 2004 [23]. The setup, see figure 6, consisted of a pumped SPDC source and a Bell state analyzer which could distinguish two Bell states. In addition, the outcome of the Bell measurement was forwarded to allow Bob to apply the correct unitary operation. The feed-forward mechanism is an essential part for future realization of quantum repeaters.

An optical fiber, installed in the sewage system underneath the river Danube was used to send one photon to Bob. The quantum channel was exposed to temperature fluc-

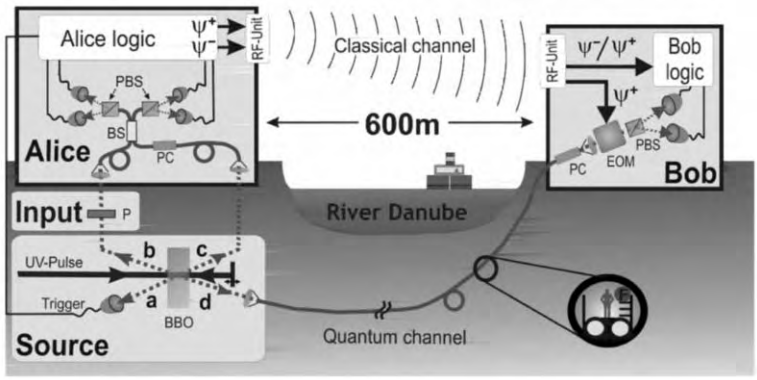


Figure 6. Teleportation under the river Danube (from [23]). Alice and Bob are located on opposite banks of the river. The quantum channel runs through the sewage system to Bob. Alice sends the outcomes of her Bell measurement via an microwave link to Bob. The classical signal arrives in time for Bob to apply to correct unitary transformation for a successful teleportation.

tuations and other environmental factors to provide an realistic test bed for the teleportation. If Alice measured successfully one of the two useful Bell states ($|\psi^-\rangle$, $|\psi^+\rangle$), she transmitted her result via a microwave link to Bob, who was situated on the other bank of the river. Due to the fact that the speed of light is only two thirds in optical fibers, then it is in air, Bob received Alice's signal 1.5μs before the arrival of his photon. This was enough time for Bob to actively set the unitary transformation necessary to obtain the correct state. The observed visibilities (85%) for a complete set of polarizations proved the operational efficiency of the system. Data was taken over several days without the necessity to readjust the polarization controllers.

3.4. Entanglement Swapping

Teleportation does not only work for unknown, but well prepared, states but also for intrinsically undefined states, like the single particle state of an entangled pair, which is completely unpolarized. This allows to teleport the entanglement of a state (entanglement swapping) [27]. Consider two sources producing the following entangled pairs: $|\psi^-\rangle_{12}$ and $|\psi^-\rangle_{34}$.

The total state of the system is then: $|\psi\rangle_{1234} = |\psi^-\rangle_{12} \otimes |\psi^-\rangle_{34}$ If a Bell measurement is performed on particles 2 and 3, then particle 1 and 4 become entangled although they never interacted.

If the state $|\psi\rangle_{1234}$ is written in the Bell basis of particle 2 and 3 we obtain [28]:

$$
\begin{aligned}
&|\psi\rangle_{1234} = \\
&|\psi^+\rangle_{14}|\psi^+\rangle_{23} - |\psi^-\rangle_{14}|\psi^-\rangle_{23} \\
&-|\phi^+\rangle_{14}|\phi^+\rangle_{23} + |\phi^-\rangle_{14}|\phi^-\rangle_{23}
\end{aligned} \tag{12}
$$

From this one can see that particles 1 and 4 always end up in an maximally entangled state. If state $|\psi^-\rangle_{23}$ is measured, then the entanglement is swapped from 1,2 to 1,4.

Entanglement swapping was experimentally shown in 2001 [25] with polarization entangled photons, in a setup similar to a normal teleportation experiment. The main

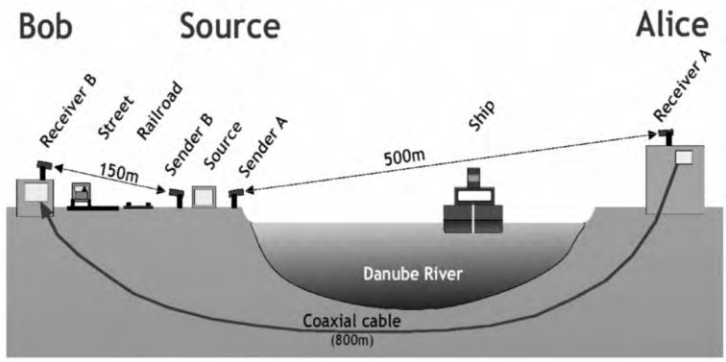


Figure 7. Entanglement distribution over the river Danube (from [35]). The SPDC source and a receiver telescope are located on one river bank, whereas the other telescope is on the far bank. Classical communication takes place through a cable running under the river bed.

difference being that the fourth photon, which was just used as a trigger before, becomes part of an entangled pair and its polarization has to be recorded. The results showed a visibility of 80% for the newly entangled state. In a further experiment [26], the entangled pair was used to show a violation of the Bell inequality ($S = 2.6 \pm 0.1$).

Entanglement swapping could in future be used to transmit entanglement over large distances with only minimal losses. Using quantum repeaters, the state would be swapped from one station to the other, along the whole chain [29]. However for practical purposes such repeaters also require quantum purification [24] and a quantum memory, which are currently investigated but are not yet sophisticated enough to be used in a practical quantum repeater scheme.

3.5. Free Space Entanglement Distribution

The direct distribution of entanglement is a crucial part in long-distance quantum information experiments. Due to the attenuation in optical fibers, the distances are limited in the order of 100 km without quantum repeaters. Long distances are only possible using a free space link. Ultimately, a free space link could reach between a satellite in orbit and one or more ground stations on earth, thus overcoming the distance limits of purely earth-based links. The following experiments can be seen as important steps towards the realization of an free-space Earth-satellite link.

3.5.1. Free Space I

An experiment distributing entangled photons with optical telescopes in free space was performed in 2003 in Vienna [35]. The two receivers Alice and Bob, located at different banks on the river Danube and with no line of sight between them, were separated by 500 m and 150 m respectively from a SPDC source, as seen in figure 7. Each photon from the entangled pair was coupled into a telescope (5 cm diameter lens) and transmitted to their respective receiver station which consisted of similar telescopes capable of polarization analysis coupled to APDs.

The detection events at Alice were transmitted back to Bob over a classical telecommunications link underneath the river[2]. Coincidence rates were analyzed and stored at Bob's station.

During the measurement, which took place at temperatures near freezing and in gusty winds, a link efficiency of 6% was achieved in each arm. A maximum coincidence rate of 15 s^{-1} could be observed. Background rates were almost entirely due to the detector dark counts, other environmental stray light was suppressed by the high spatial selectivity of the fiber coupled telescopes. Nevertheless only night time operation was possible since the onset of sun rise did still constitute an obstacle. The link attenuation of 12 dB corresponds to a value which might be achieved with state-of-the-art space technology when establishing a free-space link between an Earth-based receiver and a satellite-based transmitter.

Two methods were used to characterize the quality of entanglement after the transmission. The first parameter measured was the fidelity F of the entangled state. An average F of $87 \pm 3\%$ was obtained. Accordingly, the Bell parameter of the CHSH inequality was measured and found to be 2.41 ± 0.10, thereby demonstrating the nonlocal correlations present in the quantum state after transmission.

3.5.2. Free Space II

Encouraged by the success of the free space transmission, a more ambitious project was started in 2004. In this experiment the receivers were 8 km away from the SPDC source. The distance is more than the atmospheric equivalent, implying that more air is encountered along these 8 km than in a path going straight from the ground into space. The experiment can be seen as a test for future satellite based schemes. In addition, the free space links span between an observatory and two office buildings in Vienna, testing also the feasibility to send photons over a noisy city environment, which is important for inner city applications like free space quantum cryptography.

The telescopes used were now bigger (15 cm) to collect more light. It was found however, that after 8 km the beam already expands to 25 cm and also shows a beam wander of the same size due to atmospheric turbulence. The link efficiency was therefore only about 1.4% on average. This number proved to be too low for a distribution over two links. One could however still transmit and detect single photons via the free space link. In such a scenario the violation of the Bell's inequality can still be used to quantify the fidelity of the transmitted polarization qubit. The data was taken with one free space link, the second photon being detected directly at the source. An average coincidence rate of 84 s^{-1} was observed during the experiment. The measured photon pairs showed a violation of the Bell inequality by 14 standard deviations ($S = 2.27 \pm 0.02$).

An interesting point is the synchronization scheme. Instead of relying on a time-stable channel, time stamps were used to identify the coincidences. The time stamping was stabilized using Rubidium clocks. A software compared the times of the individual events and searched for a maximum correlation between them. The advantage of this method is that any communication channel can be used to obtain coincidence counts. This seems very practical solution since a free space link might not offer the infrastructure needed for a dedicated reference pulse.

[2]It should be noted that in this scheme, the quantum channel is above the Danube and the classical channel runs below it which is just the opposite as in the long distance teleportation experiment.

3.5.3. More Free Space

In a similar experiment, a group in Hefei succeed to distribute entanglement to two receivers each separated from the source by around 6 km [34]. Due to larger collection apertures, each of their telescopes weighing 800 kg, the link efficiencies were sufficiently high to detect 60 s^{-1} photon pairs at the distant receivers. The quality of entanglement was also sufficient to obtain a Bell parameter of ($S = 2.45 \pm 0.1$).

3.6. Multipartite Entanglement

So far we only considered entanglement between particle pairs. Multipartite entanglement also plays an important role. Examples are the GHZ [36] and W [37] state.

Here we concentrate on another class of multiparticle entangled states, the cluster state, and its role in quantum computation.

3.7. One-Way Quantum Computer

A very promising application of multipartite entanglement is the one-way quantum computer [38]. In standard measurement-based quantum computation a set of qubits is subjected to a series of one- and two-qubit gates. A register of input qubits goes through these unitary gates which implement the algorithm and ends up in an output registry where they are measured. The great technical difficulty is the realization of the two qubit gates. On one hand, the qubit should be decoupled from the environment (long decoherence time) on the other it should readily couple to a second qubit to implement a control gate.

The one-way quantum computer works in an entirely different way. The starting point of the calculation is a highly entangled multiparticle state, the so-called cluster state. After preparing the cluster, the computation proceeds via projective single qubit measurements, hence the name one-way computer given that the measurements are clearly not time reversible. The algorithm is performed via the order of measurements.

3.7.1. Preparing the Cluster

The cluster is an array of qubits which are entangled by Ising-like nearest neighbor interactions. The individual qubits are prepared in a superposition state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, where $|0\rangle$ and $|1\rangle$ are the computational basis of the qubits. A C-phase gate $|j\rangle|k\rangle \rightarrow (-1)^{jk}|j\rangle|k\rangle$ is applied between neighboring qubits to create the cluster entanglement. The choice which qubits are entangled together determines the range of algorithms the cluster is able to perform. The order in which the C-phase gates are applied does not affect the computation.

3.7.2. Computation on the Cluster

In general, one can picture a cluster computation as seen in figure 8: Input qubits are encoded in the cluster either during preparation (note that the information is distributed over the whole cluster and not in single qubits). The more qubit the algorithm processes the more rows are needed in the cluster. Computational measurements are performed column after column, from left to right, on individual qubits. The more steps an algorithm needs, the more columns of qubits are in the cluster. The output of the cluster

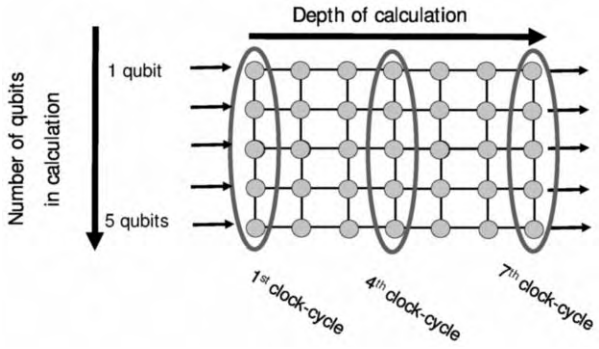


Figure 8. Schematic representation of a cluster state used in quantum computing. The number of columns determines how many qubits are processed. At each step in the algorithm a new measurement is done on a new row. The number of rows represents the number of steps in the algorithm.

computation is found in the last column. The information is not carried by the individual qubits like in a "classical" quantum computer but is processed to the end of the cluster by non-local correlations between the qubits.

Let us consider a cluster which has the input state $|\psi_{in}\rangle$ encoded in it. A single qubit operation is performed by projecting the first qubit of the cluster into one of two basis states $|\pm\alpha\rangle = (|0\rangle \pm e^{i\alpha}|1\rangle)/\sqrt{2}$. If the state is measured in $|+\alpha\rangle$ then the second qubit of the cluster (output qubit) is found in state $|\psi_{out}\rangle$, where $|\psi_{out}\rangle = HR_z(-\alpha)|\psi_{in}\rangle$. H is the Hadamard operation and $R_z(\alpha)$ is a single qubit rotation by an angle α.

A two-qubit cluster creates an output state on the second qubit which is equal to the state of the input qubit with a Hadamard gate and a rotation acted upon(the amount is determined by the measurement basis). To have an arbitrary single qubit rotation as needed in quantum computers, a four qubit cluster is needed, as shown in figure 9. Three measurement on the first three qubits with different angles on each will transfer the input state to any desirable position on the Bolch's sphere. The output state is then found on the last qubit.

Besides single-qubit rotations, two-qubit gates are also needed for universal quantum computation. However, the cluster has already C-phase gates wired into its structure. Together with the single qubit rotations C-phase gates form an universal set and are sufficient to implement any quantum algorithm. As an example we consider the cluster in figure 10. Measurements on qubits 2 and 3 are equivalent to a C-phase operation and rotations on two qubits. The logical output is found in qubits 1 and 4. A complex cluster can be thought to be made up of these elementary blocks. It should be noted that the two cluster states in figure 9 and 10 are physically the same, the only difference being the sequence of measurements on the qubits which determines the algorithm.

In essence, the entanglement within a row of the cluster is used to perform single qubit operations, whereas the entanglement between different rows serves for two qubit operations.

A four-qubit cluster state was recently experimentally realized with polarization entangled photon pairs [39]. The cluster state showed a fidelity of 63%. A universal set consisting of single-qubit rotations and two qubit operations was achieved. The circuit in figure 10 was used to produce entanglement between two, initially non interacting,

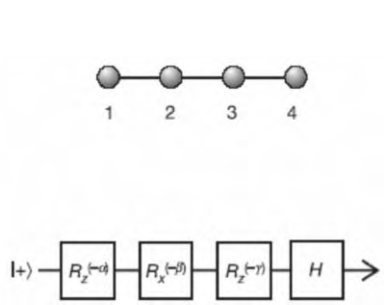


Figure 9. A four qubit cluster and its quantum circuit (from [39]). The linear sequence of measurements implements an arbitrary single qubit rotation

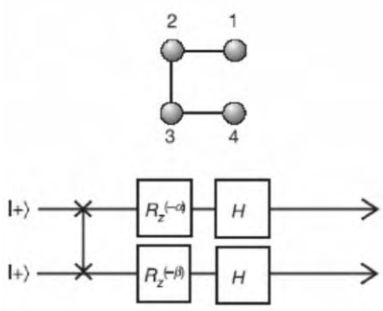


Figure 10. A four qubit cluster with measurements done on particle 2 and 3. This operation entangles qubits 1 and 4.

qubits (1 and 4). The observed entangled state showed a fidelity of 0.84 and violated the CHSH inequality with $S = 2.47 \pm 0.08$.

The four-qubit cluster state was also used to implement Grover's search algorithm [40], which is one of the best known quantum algorithms. It can identify an element in an unsorted list of N elements with only $\sqrt{O(N)}$ steps, whereas the classical counterpart needs $O(N)$ calculations. With the cluster, a specifically labeled outcome of a two-bit function could be identified among the four possible outcomes for this function (00, 01, 10, 11) with an accuracy of 90%.

The proof of principle for cluster state computation has been experimentally realized and opens the door for future implementations of this novel approach in quantum computing.

4. Quantum Key Distribution

Quantum key distribution (QKD) is probably the most technically advanced application in quantum information. It is based on the ability to detect any intruder who tries to eavesdrop on a communication link, a task which is impossible in the classical world. Due to the potential value of such systems, a great deal of effort has been put into their development, resulting in a variety of working systems. This section describes the physics behind those experimental realization and discusses in particular an experiment with entangled photon pairs.

4.1. QKD Protocols

The first security application of quantum mechanics was proposed by Wiesner [41]. He envisaged bank notes storing single photons, polarized in a secret sequence. Only the legitimate owner knows the sequence and can check if the note is correct. Any attempt to counterfeit it would fail on the grounds that the polarization state cannot be determined unambiguously without knowing the basis it was prepared in.

4.1.1. BB84

This protocol proposed by Bennet and Brassard [45] still serves as the basis for most QKD implementation. It is also very intuitive and serves as a good example to explain the physics behind QKD.

The basic argument runs as follows: Alice can prepare a photon in one of four linear polarization states: H for horizontal, V for vertical, + for 45° and - for 135° polarization. H/V and +/- form conjugate basis. Within one basis set the classical binary information is associated with one polarization, e.g. classical "0" is V or + in the two bases. If Alice wants to transmit "0" she chooses randomly between V or + to prepare her photon. The photon is transmitted to Bob, who decides in which basis (H/V or +/-) to measure it. If he chooses the same basis as Alice, his measurement determines the original quantum state and hence recovers the classical bit. If he measures in the wrong basis his measurement will reveal no information. In order to have only correlated measurements Alice and Bob must compare the basis in which the respective photon was prepared and measured. This basis reconciliation or key sifting is done publicly, but does not compromise the security of the key since no measurement results are announced.

Eve, the potential eavesdropper, can only obtain information from the transmitted photon. But here quantum mechanics comes into play and guarantees a secure transmission of the key:

- Eve **cannot copy the quantum state** of the photon and then wait for basis reconciliation to measure in the appropriate basis. The non cloning theorem [62] in quantum mechanics prohibits such actions.
- If Eve decides to measure and resend the photon, she will only half of the time pick the right basis. In the other cases Eve will get random information and the photon she sends to Bob is necessarily also uncorrelated to the photon prepared by Alice. Eve therefore introduces an error between otherwise perfectly correlated events and she can be detected. Using the **intercept-resend strategy**, Eve's tampering will amount to 25% error. More elaborate attacks [46] can reduce the error bound to 11%, above which no secure communication is any longer possible.
- Eve can exploit another weakness if the source emits states of light which contain more than one photon as is the case in the faint pulse schemes. In the so called **photon number splitting attack** (PNS), Eve keeps one photon and waits until the basis reconciliation to know in which basis she has to measure and in doing so obtains the information. Eve can also block all one-photon pulses and only let multi-photon pulses trough to Bob. In this way she can obtain the full information of the whole key.
- Eve can impersonate the legitimate sender and receiver of the key and thus gain full information. In order to prevent the **man-in-the-middle attack**, any messages exchanged between Alice and Bob must be authenticated.

4.1.2. EPR Protocol

In this protocol [47], entanglement is exploited to secure information. The entangled particles are send to Alice and Bob respectively. Now, both parties need to perform a measurement similar to Bob's in the BB84 protocol. If Alice and Bob measure in the same basis, their outcomes will always be correlated, enabling them to establish a key. To test the security of the channel, measurements are also performed in different basis

to determine the Bell parameter. If the Bell inequality is violated then the source really emits entangled pairs which have not been tampered with by Eve.

Instead of determining the Bell parameter, the EPR pairs can be implemented directly in the BB84 protocol [48]. In the original BB84 Alice prepares the state which is sent to Bob actively. In the EPR version, Alice passively prepares Bob's state by measuring her qubit and collapsing the state.

It is also worth noting that before measurement no information is contained in the entangled pair and there is therefore no information for the eavesdropper.

There are many more protocols, like the 2-state [49], 6 state [50] and the Singapore protocol [51]. Variations on the BB84 protocol include the SARG [52] protocol and Decoy states [53].

4.2. From the raw to the secret key

The transmission and detection of the physical qubits is just the first step. The raw key needs to undergo some very important stages first, before a usable secret key can be distilled.

1. Key Sifting: Alice and Bob compare the basis in which their joint measurement took place. If they measure in conjugate bases the result will be discarded. Only measurement outcomes in the same basis are kept and used further.
2. Error correction: the sifted key contains errors from a potential eavesdropper and/or imperfections in the setup. To rectify those errors, classical error correction protocols are used. During such a protocol the information exchanged is in the public domain and increases the mutual information of the eavesdropper (I(A,E)) [54]. I(A,E) is the information shared between Alice and Eve but not with Bob.
3. Privacy amplification: in this stage the information gathered by the eavesdropper, either through direct tapping of the quantum channel or by listening to the disclosed bits during error correction, is reduced to an arbitrarily small amount. However, this only works if the mutual information between Alice and Bob is bigger than I(A,E).
4. Authentication: each message which is transmitted between Alice and Bob during the QKD protocol needs to be authenticated. After each key exchange a small portion of the key is kept to authenticate future messages. Of course, in order to authenticate the very first communication, the parties must start of with a small pre-established key. Quantum key distribution becomes quantum key growing.

4.3. Physical Realization of QKD

In QKD, the physical carrier of the qubit is the photon. It is a natural choice since photons have a long decoherence time and can be easily guided over long distances. The qubit itself is represented either by the polarization state of the photon or its relative phase. Since there are no single photon sources yet, the single photon state has to be approximated by attenuated laser pulses. Entanglement based sources have the advantage that the strong correlation between the photons of a pair can be used like single photons.

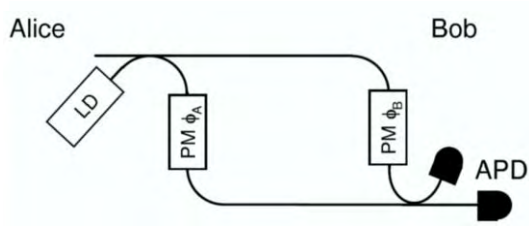


Figure 11. Single Mach-Zehnder interferometer for weak coherent pulse QKD. Alice encodes the information using her phase modulators. Bob makes the basis choice using his modulator.

4.3.1. Polarization Coding with Weak Coherent Pulses

Coherent laser pulses are only an approximation to single photons and some pulses contain more than one photon. In order to be secure against the PNS attack the mean photon number per pulse must be well below unity and must be even lowered further if, due to longer distances, the transmission loss increases [55].

Following straight from the BB84 protocol, polarization encoding is the obvious choice. It was also first demonstrated in 1992 by Bennet et al. [56] over a distance of 30 cm. Since early day fibers could not transmit the polarization states over long distances, free space versions were investigated. Among these is the transmission over 23 km between two mountains in the German Alps [57] and a 10km QKD transmission during daylight [12].

Inner city links via free space are also currently under development [58].

4.3.2. Phase Coding with Weak Coherent Pulses

As described in the previous section the polarization state of photons suffers from decohering effects in optical fibers. For this reason phase encoding is a popular choice in conjunction with fiber transmission.

In the simplest case the quantum link would simply be a very large Mach-Zehnder interferometer spanning between Alice and Bob, as introduced by Bennet [49], seen in figure 11. Alice sends attenuated laser pulses into the interferometer and encodes each pulse with a certain phase ϕ_A. In the BB84 implementation the phases 0, π encode the classical bit values 0,1. A second basis is introduced with phases of $\pi/2$, $3\pi/2$. Bob chooses his phase ϕ_B to be 0 or $\pi/2$ according to in which basis he wants to measure. If the same basis is used (i.e. the total phase change in both arms is 0 or π), the detection probabilities behind the output ports will be correlated as a function of the total phase change.

For a total phase of 0, the photons will always be detected in one APD and always in the other APD if the total phase is π. In this way a raw key can be established. The obvious limitation of this scheme is of course the requirement of phase stability over the whole interferometer during key exchange. Since in long distance applications both arms would be tens of kilometers long, it is practically impossible to keep the lengths, and hence the phase, stable to a fraction of the wavelength.

To become insensitive to the fluctuation of the fiber, Bennet suggested the use of two unbalanced Mach-Zehnder interferometers (the path length difference is longer than the coherence length of the photon), as seen in figure 12. Due to the imbalance, the photons measured by Bob can fall into one of three different arrival slots: a) short: the photon

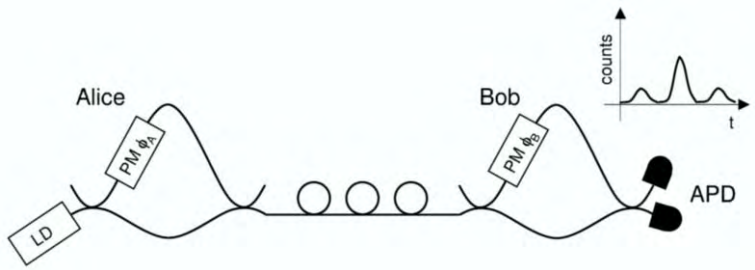


Figure 12. Unbalanced double Mach-Zehnder implementation for QKD. The inset represents the detection events as a function of time since the emission. Only counts in the center peak show correlation.

travels through both short arms of the interferometers; b) long: the photon travels through both long arms of the interferometers; c) center: The photon travels the long path in one interferometer and the short one in the other. If the two possible routes (long or short in the first interferometer) are indistinguishable, interference will be seen at Bobs output. Since a) and b) do not interfere, only the events in the center peak are counted. The protocol is then the same as in the case for the single Mach-Zehnder interferometer with Alice and Bob choosing randomly ϕ_A and ϕ_B to produce the raw key.

The advantage of the setup is that fluctuations experienced over the long single fiber transmission are the same for both interferometers (due to the indistinguishability of the routes). It still remains however to ensure that the relative imbalancement (or phase) between the two interferometers are kept stable to a fraction of the wavelength.

Such a system was used in 2004 [60] to perform a QKD experiment over more than 100 km. At 123 km the reported bit rate was around 10 Hz with a QBER of 8.9%. A system with active control of the phase difference is currently under development.

4.3.3. Plug and Play

In this scheme only one interferometer is used to avoid the problematic of keeping the relative phase stable. The setup is depicted in figure 13 and works as follows:

- Bob sends a strong laser pulse to Alice. The pulse goes through Bob's interferometer and splits into P1 (short arm) and P2 (long arm). No phase modulation is applied at this moment and both pulses, separated by about 200 ns, leave the PBS towards Alice.
- On Alice's side, part of P1 is used to provide a timing reference. The rest of P1 and P2 are incident on a Faraday mirror. The effect of the mirror is to transform an incoming polarization state into its orthogonal counterpart. Alice then applies a phase shift on P1 in order to encode the information. Before leaving Alice, P1 and P2 are attenuated to single photon level.
- The photons are now traveling back to Bob and arrive there orthogonally polarized. P1 is now reflected and travels the long arm in Bob's interferometer whereas P2 takes the short arm. Bob makes his basis choice using ϕ_B. P1 and P2 interfere at the coupler and the output port chosen is again a function of ϕ_A and ϕ_B.

More details on this system can be found in [59]. Due to the initially strong pulses the system is liable to suffer from backscattering of the fiber. At high transmission rates the strong pulses will intersect the single photon pulses traveling in the opposite direction.

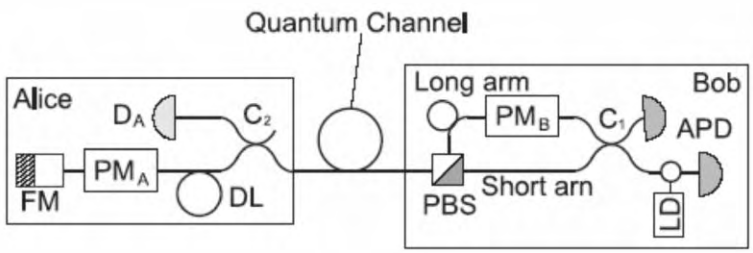


Figure 13. Scheme of a plug&play setup in QKD (from [59]). Bob's side consists of a laser diode, an interferometer and phase modulator. The station at Alice is made of a Faraday mirror, optical attenuator and a phase modulator

Even if the fraction of backscattered light is very low, it suffices to wash out the single photon signals and to introduce errors. One way to overcome this is the operation in burst mode where a train of pulses is sent. Alice uses a storage line to delay the single photons until all the strong pulses have left the fiber. Only then are the weak signals transmitted back to Bob. Although this completely solves the problem, the price to pay is a lower transmission rate.

A field test over 67 km in 2002 [61] showed a raw key rate of 150 Hz. The QBER was around 6%. The system consisted of two 19 inch racks and the transmission took place over standard telecom fibers.

4.3.4. Phase Coding With Entangled Photon Pairs

This class of experiments relies on the correlations present in entangled photon pairs. Entangled pairs have the advantage not to be liable to PNS attacks. Even if a double pair is created by the source the correlations are strictly limited to a single pair. No information exists between photons of different pairs.

In the phase coding scheme, as seen in figure 14, which is very similar to the double Mach-Zehnder setup of the previous paragraph, a SPDC source is used to produce photon pairs. If the pump laser has a large coherence length then the time at which the photon pair is produced becomes uncertain, even though the two photons are still emitted simultaneously.

There are 3 possible detection times at Bob's side with respect to Alice's measurement (see inset fig. 14): The first and last peak corresponds to events where the photons either went the short/long path or long/short respectively at Alice/Bob, so these counts are discarded. Events where both photons take the long or the short arm result in counts in the middle peak. If the coherence length of the pump laser is larger than the pathlength difference in the interferometers, one observes two photon interference as a function of the relative phases ϕ_A and ϕ_B [42].

4.3.5. Polarization Coding With Entangled Photon Pairs

This scheme is similar to the original BB84 proposal. But instead of sending a photon from Alice to Bob, a source emitting polarization entangled pairs of photons is used. The pair is split and the photons are sent to Alice and Bob respectively. On each side a BB84 module randomly projects the photons in one of the two bases (H/V or +/-). If the same basis is chosen the polarization measurements will be correlated.

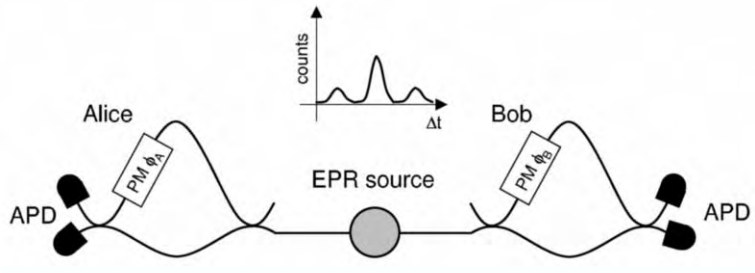


Figure 14. QKD system using energy-time entangled photons. A SPDC source emits a photon pair into two unbalanced Mach-Zehnder interferometers. Detection events in coincidence (central peak) yield correlated outcomes.

The first experiment with polarization entangled photons was performed in 1998 [63]. Photons at 810 nm were distributed over 400 m and a key was generated at the rate of 1 kbit/s with an error rate of 3%. The key was further used as a one-time pad to send encrypted data.

4.4. A Real World Experiment

In April 2004, a bank transfer, secured by a quantum key, was successfully accomplished [43]. The aim of this experiment was to perform the first experiment which implements all the necessary steps of a full QKD protocol and to test the feasibility of possible future applications. The trial should not happen under laboratory conditions but should take place in a real-world scenario. The choice of location fell to the Vienna City Hall and the headquarters of a national bank.

4.4.1. Experimental Setup

The setup used is shown in figure 15. On Alice's side (bank) it consisted of a compact source of entangle photons, a BB84 detection unit and QKD-hardware. A type-II SPDC setup was used to produce entangled photon pairs in the $|\psi^-\rangle$ state with a central wavelength of 810 nm and a FWHM bandwidth of 5.6 nm. With a 16 mW pump laser about 8200 pairs could be detected locally. One photon was send directly to Alice's detection module whereas the other photon was send to Bob via a 1.45 km fiber link. At both detection units, the incoming photon chooses randomly between the H/V and +/- basis. In this way the BB84 protocol was directly implemented.

The fiber used for the quantum channel was single mode for 810 nm and showed an overall attenuation of 6 dB. Beside the quantum channel, 3 additional standard telecom fibers were used to handle the synchronization pulses and classical data transfer. The whole fiber bundle was installed in the Vienna sewage system, where it was subjected to environmental influences like temperature fluctuations.

The polarization rotation induced by the fiber was compensated using polarization controllers at the source. Once the controllers were adjusted, the polarization proved to be stable for hours and no active compensation was necessary during key exchange. For longer operation an active control will be necessary but only for long time scales (few hours).

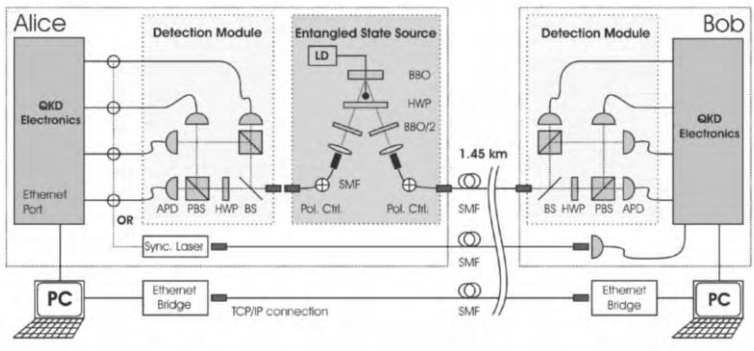


Figure 15. Sketch of the polarization entangled QKD setup (from [43]). A SPDC source, located at Alice, produces entangled photon pairs. One photon is directly analyzed at Alice. The other photon is transmitted to Bob over a 1.45 km single mode fiber. Once a photon is measured at Alice, an optical trigger pulse is sent to Bob to establish a time reference. The detector events and synchronization pulses are fed on both sides into an electronic hardware which performs all necessary QKD operations to yield the secure key.

The idea of the QKD-hardware is to have all the steps required for a full automated QKD exchange on an embedded electronics board which is compatible with conventional telecommunication standards. Moreover, the electronics was recently extended not only to handle the secret key as a resource for a one-time pad, but also to symmetrically expand the key to increase the encrypted data rate (with slight reduction in the security of the key).

4.4.2. Results

During a 18 minute long key exchange a total of 250000 bits of sifted key was recorded. To implement the classical algorithms the data was divided into blocks of 2500 bits each. 25% of the bits in each block were publicly compared to estimate the QBER. An average error value of 6.3% was found in this way.

Classical error correction using CASCADE [44] reduced the key further to about 110000 bits. It was seen that the number of disclosed bits by CASCADE was only weakly dependent on the estimated error rate. On the other hand reduction in key during privacy amplification is strongly correlated with the QBER and shows large fluctuations due to the strong statistical deviations of the QBER estimated from only small sample sizes. The final secure quantum key after privacy amplification amounted to 79000 bits. This corresponds to an average secure key of 76 bits/s.

The experiment demonstrated the feasibility of entanglement based quantum cryptography. Furthermore it showed the state of QKD where all the necessary calculation are done in real time. It also proved that such a system can be operated outside the protective laboratory, namely in a real world scenario with strong environmental exposure.

Summary & Outlook

Entanglement serves as an important ingredient in quantum information applications. Several methods of production of entangled photons were presented. Depending on the type of application, the transmission of entangled photons is either via optical fibers or

in free-space. The former has the advantage that the photons can be easily directed to the recipient but dispersion effects lead to decoherence. Free space links are currently developed with the aim to offer long distance transmission of quantum states, ultimately via satellite. Repeater stations would allow the distribution of entanglement on the ground. The first steps towards such schemes are under way, including long distance teleportation. Highly entangled multiparticle states might be a way to tackle quantum computation from an altogether different perspective. A four-photon cluster state was used as proof of principle for one-way quantum computation. Quantum cryptography shows very neatly the advantages of using the rules of quantum mechanics as opposed to classical information.

The first steps in the field were on a more theoretical nature and only a handful of experimental data. However, the technological advances in the past years made it possible to realize many of the past dreams in the laboratory. Successful implementation of early experiments are on the way to become everyday technology (e.g. detectors, random number generators). Quantum cryptography seems to be the next application to be leaving the laboratory and be implemented in commercial products. Given the relatively recent origins of the field one can expect still many more exciting discoveries and applications to come.

Acknowledgements

We would like to thank Andreas Poppe, Markus Aspelmeyer, Philip Walther and Rupert Ursin for valuable discussion and for providing material.

References

[1] P. Kwiat et al., *Phys. Rev. Lett.* **75** (1995), 4337
[2] M.A. Horne, A. Shimony and A. Zeilinger, *Phys. Rev. Lett.* **62** (1989), 2209
[3] J. Rarity et al., *Phys. Rev. Lett.* **64** (1990), 2495
[4] J.D. Franson, *Phys. Rev. A* **44** (1991), 4552
[5] P. Kwiat et al., *Phys. Rev. A* **47** (1993), 2472
[6] P. Kwiat et al., *Phys. Rev. A* **60** (1998), 773
[7] Fiorentino et al., *Optics Express* **13** (2005), 127
[8] A. Yoshizawa and H. Tsuchida, *App. Phys. Lett.* **85** (2004), 2457
[9] Ribordy et al., *Phys. Rev. A* **63** (2000), 012309
[10] König et al., *submitted to Phys. Rev. A* (2004) quant-ph/0409162
[11] M. Pelton et al., *Optics Express* **12** (2005), 3573
[12] R.J. Hughes et al., *New J. Phys.* **4** (2002), 43
[13] C. Langrock et al., *Optics Letters* **30** (2005), 1725
[14] G. Brida et al., *J. Mod. Opt.* **50** (2003), 1757
[15] D. Bohm, *Quantum Theory*, Prentic-Hall, Enlewood Cliffs (1951), 614
[16] A. Einstein, B. Podolsky. N. Rosen, *Phys. Rev.* **47** (1935), 777
[17] J.S. Bell *Speakable and Unspeakable in Quantum mechanics*, Cambridge U.P., Cambridge (1987)
[18] J.F. Clauser et al., *Phys. Rev. Lett.* **23** (1969), 880
[19] P. Grangier, *Nature* **409** (2001), 774
[20] G. Weihs et al., *Phys. Rev. Lett.* **81** (1998), 5039

[21] S.J. Freedman and J.F. Clauser, *Phys. Rev. Lett.* **28** (1972), 938
[22] A. Aspect et al., *Phys. Rev. Lett.* **49** (1982), 1804
[23] R. Ursin et al., *Nature* **430** (2004), 849
[24] J.W. Pan et al., *Nature* **423** (2003), 6938
[25] J.W. Pan et al., *Phys. Rev. Lett.* **86** (2001), 4435
[26] T. Jennewein et al., *Phys. Rev. Lett.* **88** (2002), 17903
[27] M. Zukowski et al., *Phys. Rev. Lett.* **71** (1993), 4287
[28] D. Bouwmeester, A. Ekert and A. Zeilinger (Eds.): *The Physics of Quantum Information*, Springer, Berlin (2001)
[29] H.J. Briegel et al., *Phys. Rev. Lett.* **81** (1998), 5932
[30] D. Bouwmeester et al. *Nature* **390** (1997), 575.
[31] C.K. Houng et al., *Phys. Rev. Lett.* **59** (1987), 2044
[32] N. Lütkenhaus et al., *Phys. Rev. A* **59** (1999), 3295
[33] C.H. Bennet et al., *Phys. Rev. Lett.* **70** (1993), 1895
[34] C. Z. Peng et al., *Phys. Rev. Lett.* **94** (2005), 150501
[35] M. Aspelmeyer et al., *Science* **301** (2003), 621
[36] D.M. Greenberger, M.A. Horne and A. Zeilinger , *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, Kluwer Academics, The Netherlands 1989, 73–76
[37] W. Dür et al., *Phys. Rev. A* **62** (2000), 062314
[38] R. Raussendorf and H.J. Briegel, *Phys. Rev. Lett.* **86** (2001), 5188
[39] P. Walther et al., *Nature* **434** (2005), 169
[40] L.K. Grover *Phys. Rev. Lett.* **79** (1997), 325
[41] S. Wiesner, "Conjugate Coding", written circa 1970, published in *Sigact News* **15** (1983), 78
[42] A.K. Ekert et al.: Practical Quantum Cryptography Based on Two-Photon Interferometry, *Phys. Rev. Lett.* **69** (1992), 1293–1295.
[43] A. Poppe et al., *Optics Express* **12** (2004), 3865
[44] G. Brassard and L. Salvail: Lecture Notes in Computer Science, Springer, **765**, 410–423
[45] C.H. Bennet and G. Brassard, *Proceeding of Int. Conf. Computers, Systems & Signal Processing, Bangalore, India* **175** (1984)
[46] P.W. Shor and J. Preskill, *Phys. Rev. Lett.* **85** (2000), 441
[47] A. Ekert, *Phys. Rev. Lett.* **67** (1991), 667
[48] C.H. Bennet et al., *Phys. Rev. Lett.* **68** (1992), 557
[49] C.H. Bennet, *Phys. Rev. Lett.* **68** (1992), 3121
[50] D. Bruss, *Phys. Rev. Lett.* **81** (1998), 2598
[51] J. Anders et al., *quant-ph/0505069* (2005)
[52] V. Scarani et al. *Phys. Rev. Lett.* **92** (2004), 057901
[53] H.K. Lo et al., *Phys. Rev. Lett.* **94** (2005), 230504
[54] M. Curty et al., *Phys. Rev. Lett.* **92** (2004), 217903
[55] N. Lütkenhaus, *Phys. Rev. A* **61** (2000), 052304
[56] C.H. Bennet at al., *J. Cryptology* **5** (1992), 3
[57] C. Kurtsiefer et al., *Nature* **419** (2002), 450
[58] T. Schmitt-Manderbach et al., LMU Munich,*Presentation at CLEO Europe* (2005)
[59] N. Gisin et al., *Rev. Mod. Phys.* **74** (2002), 145
[60] C. Gobby et al., *App. Phys. Lett.* **84** (2004), 3762
[61] D. Stucki etz al., *New J. Phys.* **4** (2002), 41.1
[62] W.K. Wootters and W.H. Zureck *Nature* **299** (1982), 802
[63] T. Jennewein et al., *Phys. Rev. Lett.* **84** (2000), 4729

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Preparation and Measurement of Qutrits Based On Single-Mode Biphotons

Sergei Kulik [a]
[a] *Faculty of Physics, Moscow M.V.Lomonosov State University, Russia*

Abstract. We have discussed an interferometric method of preparing three-level quantum optical systems that relies on the polarization properties of single-mode two-photon light generated via spontaneous parametric down-conversion. For adequate state measurement the quantum tomography procedure developed for biphotons-qutrits was applied. A specific sequence of nontrivial states was generated and measured with high fidelity. For some of the states, forming the mutually unbiased bases, we applied the operational orthogonality criterion and experimentally demonstrated the orthogonality of the selected states. This suggests the possible usage of our setup (with some improvements) in a practical realization of a QKD protocol.

1. Introduction

The ability to generate, transmit and measure quantum systems is greatly desired in the emerging field of quantum information technology. A large number of protocols that rely on the properties of two-level quantum systems have been proposed, and many have subsequently been realized. A general overview is provided in [1]. Naturally, the question of whether it is possible to increase the dimensionality of the systems used as information carriers has arisen, and with it interest in the new features that such an extension can offer. The simplest extension involves the use of three-level quantum systems (qutrits). In the field of quantum information, those systems that can be completely described in terms of three orthogonal states are called qutrits (quantum trits). In the case of a pure state, the wave function of the three-level system can be written as:

$$|\Psi\rangle = c_1|1\rangle + c_2|2\rangle + c_3|3\rangle \tag{1}$$

where $|1\rangle, |2\rangle, |3\rangle$ are the orthogonal basis states. The complex coefficients c_j are the amplitudes of the basis states $|j\rangle$ and satisfy the following normalizing condition

$$\sum_{i=1}^{3} |c_j|^2 = 1 \tag{2}$$

Decomposition (1) is the generalization of the definition of qubit, where the dimension of the quantum system has been increased to $d = 3$.

Determining the complex coefficients c_j is a significant question in quantum information science. Once this problem has been resolved, one can expect the realization of the a variety of information protocols and devices, such as the protocols of error correction and quantum repeaters. From a fundamental point of view, the question of the minimal set of measurements that is needed for the complete description of the system's state is also very important. In some cases it is not necessary to perform a complete set of measurements to define a system's purity [2]. Of course, for different types of quantum state, one uses different types of measurement procedures. For example, for a squeezed state of light, methods of homodyne tomography are being developed [3]. In principle, these allow the determination of the density matrix of n-photon Fock states [4]. For the polarization-squeezed [5] and scalar [6] light, the fluctuations of Stokes parameters are analyzed and the quasi-probability function is restored [7]. In the case of two-photon fields, one measures the set of the fourth order field moments in different spatial and polarization modes [8]. In context of a particular experimental procedure, the *a priori* information about the properties of the examined state plays an important role.

Recently, new quantum key distribution (QKD) protocols were proposed that specifically use qutrits [9,10] and the eavesdropping analysis showed that these systems are more robust against a specific classes of eavesdropping attacks [11,12]. The other advantage of using multilevel systems is their possible implementation in fundamental tests of quantum mechanics [13], giving more divergence from classical theory. The usage of multilevel systems also provides the possibility to introduce novel protocols, which cannot be implemented with the help of qubits such as Byzantine Agreement [14] and Quantum Coin Tossing [15].

Recent experimental realizations of qutrits have relied on several processes. In one case, an interferometric procedure is used, where the attenuated laser pulses are sent in a multi-armed interferometer [16]. The number of arms defines the dimensionality of the system. The identification of states is performed either through pulse delays (temporal basis) or by the presence of constructive interference in a certain arm of an interferometer that is put into a registration system (energy basis). Other techniques rely on the orbital angular momentum of single photons [17,18] or on postselection of qutrits from four-photon states [19]. Recently, the strong transversal correlations between spatial sub-structure of signal and idler modes of two parametric down-converted photons were used to create a multidimensional system [20,21].

2. Biphotons as Qutrits

Here we consider an experimental realization of qutrits that exploits the polarization state of a single-mode biphoton field. This field, which consists of pairs of correlated photons, is most easily obtained with the help of spontaneous parametric down-conversion (SPDC). By saying single-mode we mean that the pairs of photons that form a biphoton have equal frequencies and propagate along the

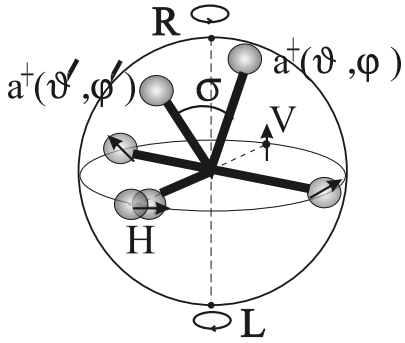


Figure 1.

same direction, so it is impossible, even in principle, to distinguishing the photons using these parameters. The pure polarization state of such a field can be written as the coherent superposition of three basic states.

$$|\Psi\rangle = c_1|2,0\rangle + c_2|1,1\rangle + c_3|0,2\rangle, \tag{3}$$

where $c_i = d_i e^{i\phi_i}$ are the complex probability amplitudes. The states $|2,0\rangle$ and $|0,2\rangle$ correspond to a type I interaction where twin photons have collinear polarization vectors (for example state $|2,0\rangle$ corresponds to two photons being in horizontal H polarization mode), and state $|1,1\rangle$ is obtained via a type II process, where photons are polarized orthogonally (say one of them is in H and the other one is in the V mode). There exists an alternative representation of state $|\Psi\rangle$ that maps the state onto the surface of the Poincaré sphere (Figure 1) [22]

$$|\Psi\rangle = \frac{a^\dagger(\vartheta,\varphi)a^\dagger(\vartheta',\varphi')|vac\rangle}{\sqrt{||a^\dagger(\vartheta,\varphi)a^\dagger(\vartheta',\varphi')|vac\rangle||}}, \tag{4}$$

where $a^\dagger(\vartheta,\varphi)$ and $a^\dagger(\vartheta',\varphi')$ are creation operators of a photon in a certain polarization mode $a^\dagger(\vartheta,\varphi) = \cos\frac{\vartheta}{2}a_H^\dagger + e^{i\phi}\sin\frac{\vartheta}{2}a_V^\dagger$ where the operators $a_{H,V}^\dagger$ indicate photon creation in the horizontal and vertical polarization modes. The parameters $\vartheta,\vartheta' \in [0,\pi]$ and $\varphi,\varphi' \in [0,2\pi]$ are polar and azimuthal angles that define the position of each photon on a surface of a sphere. The values of the angles can be calculated using the amplitudes and phases of c_i.

The goal of this article is to demonstrate the ability to prepare and measure any given polarization qutrit state $|\Psi\rangle$. As a straightforward and practical example of states to be prepared, we choose the specific sequence that was suggested in [9]. This sequence of 12 states forms four mutually unbiased bases with three states in each, and can be used in an extended version of the BB84 QKD protocol for qutrits. The 12 states are defined in Table 1.

	State	d_1	d_2	d_3	ϕ_1	ϕ_2	ϕ_3
$\|2,0\rangle$	α	1	0	0	0	0	0
$\|1,1\rangle$	β	0	1	0	0	0	0
$\|0,2\rangle$	γ	0	0	1	0	0	0
$\frac{1}{\sqrt{3}}(\|2,0\rangle + \|1,1\rangle + \|0,2\rangle$	α'	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	0	0	0
$\frac{1}{\sqrt{3}}(\|2,0\rangle + e^{2\pi i/3}\|1,1\rangle + e^{-2\pi i/3}\|0,2\rangle$	β'	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	0	$\frac{2\pi}{3}$	$-\frac{2\pi}{3}$
$\frac{1}{\sqrt{3}}(\|2,0\rangle + e^{-2\pi i/3}\|1,1\rangle + e^{2\pi i/3}\|0,2\rangle$	γ'	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	0	$-\frac{2\pi}{3}$	$\frac{2\pi}{3}$
$\frac{1}{\sqrt{3}}(e^{2\pi i/3}\|2,0\rangle + \|1,1\rangle + \|0,2\rangle$	α''	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{2\pi}{3}$	0	0
$\frac{1}{\sqrt{3}}(\|2,0\rangle + e^{2\pi i/3}\|1,1\rangle + \|0,2\rangle$	β''	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	0	$\frac{2\pi}{3}$	0
$\frac{1}{\sqrt{3}}(\|2,0\rangle + \|1,1\rangle + e^{2\pi i/3}\|0,2\rangle$	γ''	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	0	0	$\frac{2\pi}{3}$
$\frac{1}{\sqrt{3}}(e^{-2\pi i/3}\|2,0\rangle + \|1,1\rangle + \|0,2\rangle$	α'''	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$-\frac{2\pi}{3}$	0	0
$\frac{1}{\sqrt{3}}(\|2,0\rangle + e^{-2\pi i/3}\|1,1\rangle + \|0,2\rangle$	β'''	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	0	$-\frac{2\pi}{3}$	0
$\frac{1}{\sqrt{3}}(\|2,0\rangle + \|1,1\rangle + e^{-2\pi i/3}\|0,2\rangle$	γ'''	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	0	0	$-\frac{2\pi}{3}$

Table 1.

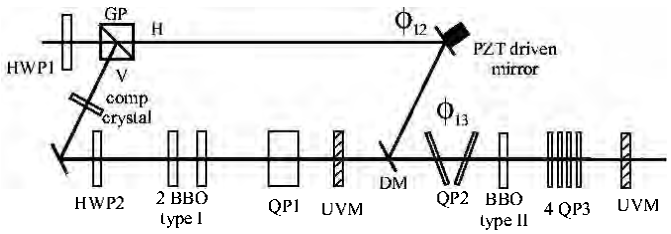


Figure 2.

The preparation part of the setup (Figure 2) is built on the base of a balanced Mach-Zehnder interferometer (MZI) [23]. The pump part consists of a frequency doubled "Coherent Mira 900" femtosecond laser, driven at central wavelength of 800 nm, 75 MHz repetition rate and with a pulse width of 100 fs. A 1 mm BBO crystal was used as a frequency doubler. As a result, we achieved an average power of 400 mW at 400 nm at the output of the crystal, but it was reduced down to 90 mW after passing through a series of pinholes and a polarizing Glan-Tompson prism (is not shown on the figure). The second Glan-Tompson prism (GP), which transmits the horizontally polarized fraction of the pump and reflects the vertically polarized fraction, serves as an input mirror of the MZI. The reflected part, after passing through the compensating BBO crystal and a half-wave plate (HWP2), pumps two consecutive 1 mm thick type I BBO crystals, which have perpendicularly oriented optical axes. The biphotons from these crystals pass through a 10 mm quartz plate (QP1) that serves as a compensator, and the pump is reflected by the UV mirror. Then the biphotons arrive at a dichroic mirror (DM) that is designed to transmit them and to reflect the horizontally polarized component of the pump that was coming from the upper arm of MZI. A piezoelectric translator (PZT) is used to change the phase shift of this component of the pump with respect to the one propagating in the lower arm. The UV

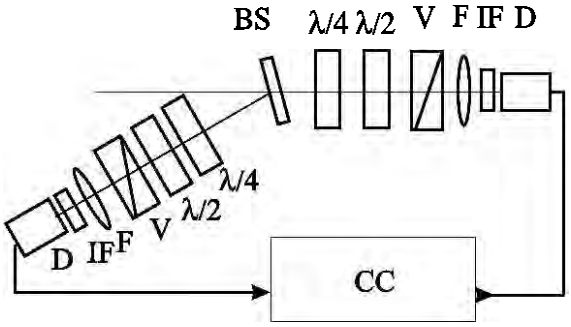


Figure 3.

beam, reflected from DM serves as a pump for a 1mm thick type II BBO crystal. Two 1 mm quartz plates (QP2) introduce a phase shift between horizontally and vertically polarized type I biphotons, and a set of four 1 mm thick quartz plates (QP3) compensate for the group velocity walk-off between orthogonaly polarized type II biphotons during their propagation in BBO crystal. The measurement setup (Figure 3) consists of a Brown-Twiss scheme with a non-polarizing 50/50 beamsplitter; each arm contains consecutively placed quarter- and half waveplates and an analyzer that is set to transmit the vertical polarization. This sequence of waveplates and analyzer is refered as a polarization filter. Interference filters of 5 nm bandwidth, centered at 800 nm and pinholes are used for spectral and spatial modal selection of biphotons. We used EGG-SPCM-AQR-15 single photon counting modules as our detectors. We assume that the source generating qutrits is stationary. Since each measurement eliminates a qutrit one has to be sure that there are a lot of copies of the initial state; each copy must be prepared in the same quantum state. Such an *ensemble approach* guarantees that the experimentalist deals with the same quantum state in all trials. In other words, the outcomes provide him with the information about the same quantum state and elimination of a particular state does not affect the rest of them.

The measurement of the generated states is performed using the tomography protocol that was developed for polarization qutrits [24,25,26,27]. What does it mean to measure the unknown state (3)? From the experimental point of view, it means that the experimentalist has to measure a complete set of real parameters (moments) determining the state. To do this, the state must be subject to a set of unitary polarization transformations and projective measurements. By doing this one gets outcomes with probabilities proportional to the corresponding field moments or their linear combination. This procedure is known as quantum tomography. The quantum state can be represented either by using a wave function, density matrix, or quasi-probability function (Wigner function). Probably the correct way to use the term "quantum tomography" is only for the reconstruction of the quasi-probability function because it gives the graphical representation of the state as a 3D plot. Nevertheless the term "quantum tomography" is also used for a general procedure of complete state reconstruction. For a selection of papers where this procedure was realized experimentally, let us mention the works [28,29,30], which utilise states defined by continuous variables. For

states characterized by discrete variables, such as two polarization-spatial qubits, quantum tomography was realized in [8]. Recently quantum tomography has been performed for orbital angular momentum entangled qutrits [31].

The physical idea behind the tomography procedure is performing measurements of an appropriate complete set of observables called quorum [32] or just "looking" at the state from different positions. The minimal number of such positions might be the number of real parameters determining the state. According to Bohr's complementarity principle, it is impossible to measure all field moments simultaneously, operating with a single qutrit only. So to perform a complete set of measurements one needs to generate many copies of a quantum state.

First of all, let us mention that at present, the only realistic way to register single-mode biphoton field is using the Brown-Twiss scheme. This scheme consists of a beam-splitter followed by a pair of detectors connected with the coincidence circuit. It means that registration of a single biphoton, which carries the state (3), can give only a single event at the output of the experimental setup with some probability. So the statistical treatment of the outcomes becomes extremely important. For correlations between polarization degrees of freedom, which is essential in the case under consideration, the Brown-Twiss scheme must be accomplished with polarization filters introduced into each arm.

The idea of the measurement method is splitting the input state into two spatial modes and performing transformations over two photons independently (Figure 3). These transformations can be performed using polarization filters placed in front of detectors. Each filter consists of a sequence of quarter- and half-wave plates and a polarization prism, which picks out definite linear polarization, for example, the vertical one. A narrow- band filter centered at the doubled pump wavelength $\lambda = 2\lambda_p$ serves to make biphotons emitted from different sources indistinguishable in frequency as well as to reduce the background noise. An event is considered to be detected if a pulse appears at the output of the coincidence circuit. In approximately half of the trials, one of the photons (signal, by convention) forming a biphoton goes to one of the detectors, while the other one (idler) goes to the other detector. In the remaining cases, both photons appear in the same output beam-splitter arm, and these events are not selected because they do not contribute to coincidences. This protocol was suggested and developed in [24,25,26,27].

A similar protocol was earlier considered in detail [8] for estimating the polarization state of a biphoton field, generated in a frequency degenerate non-collinear mode. In this case the biphoton field is represented as a pair of polarization qubits.

The configuration of our measurement setup (Figure 3) allows us to verify the orthogonality of states that belong to the same basis. According to the orthogonality criterion for biphoton polarization states [33], two polarization states $|\Psi_a\rangle$ and $|\Psi_b\rangle$ are orthogonal if one observes zero coincidence rate in the Brown-Twiss scheme, provided that state $|\Psi_a\rangle$ is on the input, and polarization filters in each arm are tuned to assure maximal transmittance of each photon that forms the state $|\Psi_b\rangle$ (set state). Experimentally, the criterion was verified in [34].

The following question arises - how many (real) parameters should be measured to characterize completely the unknown state of biphoton field? From the definition and properties of the density matrix, it follows that in the case of the

pure state, the number of real parameters that define the state of the system that has a dimension d is equal to $2d-2$, and in the case of a mixed state is equal to d^2-1 . Correspondingly for the qutrits in the first case one needs to know four real numbers, in the second case, eight. Considering that in an experiment one measures the non-normalized state's amplitudes and normalization conditions need to be checked every time after measuring all three diagonal elements of density matrix, we obtain that for pure states, five moments need to be measured and in the mixed state, nine. Before we go further into the discussion of the quantum tomography protocol, we notice that the measurement procedure always leads to the destruction of our state, caused by its interaction with the classical measuring device. So, when we speak about the input state, we always have in mind that it is introduced by a large enough set of copies, and some of them can be destroyed by the measurement. The results of the measurements will be applied to the rest of the ensemble; this procedure lies at the heart of the ensemble method of quantum measurements.

Finally, in order to reconstruct the density matrix of a measured state (that is generally mixed) one has to perform nine projective measurements of the fourth-order moments of the field for the different settings of polarization filters (Figure 3). The polarization density matrix can be defined in terms of these moments by

$$\rho = \begin{pmatrix} A/2 & D/\sqrt{2} & E/2 \\ D^\star/\sqrt{2} & C & F/\sqrt{2} \\ E^\star/2 & F^\star/\sqrt{2} & B/2 \end{pmatrix} \tag{5}$$

with

$$\begin{aligned} A &\equiv \langle a_x^{\dagger 2} a_x^2 \rangle, \quad B \equiv \langle a_y^{\dagger 2} a_y^2 \rangle, \quad C \equiv \langle a_x^\dagger a_y^\dagger a_x a_y \rangle \\ D &\equiv \langle a_x^{\dagger 2} a_x a_y \rangle, \quad E \equiv \langle a_x^{\dagger 2} a_y^2 \rangle, \quad F \equiv \langle a_x^\dagger a_y^\dagger a_y^2 \rangle \end{aligned} \tag{6}$$

Here, $a_{x,y}^\dagger$ are the creation operators of the photons in horizontal and vertical polarization modes correspondingly. It can be seen that the diagonal components of ρ are real. They characterize the intensity fluctuations in parallel (A and B) or orthogonal (C) polarization modes. Non-diagonal elements D, E, F in general case are complex. Table 2 presents the quantum tomography protocol for biphoton-qutrits. Each line in the table contains the orientation of the half ($\theta_{1,2}$) and quarter ($\chi_{1,2}$) wave plates in the measurement block and corresponds to the measured field moment (last column).

3. Compensation

In order to cause the three terms in superposition (3) to interfere, we must achieve their perfect overlap in frequency, momentum and time domains. From the experimental point of view this means that the biphoton wavepackets coming from the two type I crystals and from the type II crystal must overlap. The overlap in frequency domain is achieved by the usage of 5nm bandwidth interference filters

ν	χ_1	θ_1	χ_2	θ_2	Field moment to be measured
1.	0°	45°	0°	45°	$A/4$
2.	0°	45°	0°	0°	$C/4$
3.	0°	0°	0°	0°	$B/4$
4.	45°	0°	0°	0°	$1/8(B+C+2ImF)$
5.	45°	22.5°	0°	0°	$1/8(B+C-2ReF)$
6.	45°	22.5°	0°	45°	$1/8(A+C-2ReD)$
7.	45°	0°	0°	45°	$1/8(A+C+2ImD)$
8.	−45°	11.25°	45°	78.75°	$1/16(A+B-2ImE)$
9.	−45°	22.5°	45°	−22.5°	$1/16(A+B-2ReE)$

Table 2.

and the overlap in momentum is ensured by the usage of pinholes that select one spatial mode of biphoton field. However, the overlap in time cannot be achieved easily when using a pulsed laser source because it is necessary to compensate for all the group delays that biphoton wavepackets acquire during their propagation through optical elements of the setup [35]. We performed independent measurements of all the group delays at the pump and biphoton frequencies and for different polarization settings in a bulk Michelson interferometer. It was found that in order to overlap type I biphotons with type II, the pump pulse from the lower arm must be delayed by 50 ps with respect to the upper arm pulse. This was achieved by inserting a 2 mm BBO crystal in the lower arm. The overlap between states $|2,0\rangle$ and $|0,2\rangle$ was achieved by inserting 10 mm quartz plate directly after the two type I BBO crystals. The high visibility of interference patterns that we obtained was a criterion for a good compensation.

4. Calibration

In order to create a given qutrit state, we need to have independent control over four parameters - two relative amplitudes and two relative phases. In the experiment, we used HWP1 to control the amplitude of the state $|1,1\rangle$, and HWP2 to control the relative amplitudes of states $|2,0\rangle$ and $|0,2\rangle$. The calibration of these elements can be done by measuring moments A, B and C in the tomography setup [24,25,26,27]. For our concrete experiment we have equal weights of each basis state in the superposition, so we set our waveplates in such a way that these moments fulfill the condition: $A = B = 2C$. The relative phase φ_{13} between states $|2,0\rangle$ and $|0,2\rangle$ can be controlled with the help of rotating quartz plates (QP2). Rotation of these plates along the optical axis introduces a phase shift between the horizontally and vertically polarized biphotons. The resulting state of the biphoton field after passing through these plates can be written as $|\Psi'\rangle = |2,0\rangle + e^{i\varphi_{13}}|0,2\rangle$. It is known that for $\varphi_{13} = \pi/2$ this state transforms to $|R,L\rangle$, where R and L indicate right and left hand circular polarizations and for $\varphi_{13} = \pi$ it transforms to $|45°,-45°\rangle$. Therefore, by varying the phase φ_{13} by means of rotation the QP, one can observe an interference pattern when the polarizers in a Brown-Twiss scheme are set to transmit $+45°$ and $-45°$ polarized light. This effect, known as "space-time interference", can be used to calibrate φ_{13} since we

can assign value "0" to the position of the minimum of interference pattern, and "π" to the position of the maximum. The relation of the phase φ_{12} between the state $|\Psi'\rangle$ and $|1,1\rangle$ to the voltage applied on PZT can be found by monitoring the pump interference pattern in the M-Z interferometer.

It is important to note that in order to ensure the resulting state is close to a pure state, it is necessary to achieve as high visibility as possible in a space-time interference experiment with type I biphotons and polarization interference of type II biphotons. The obtained visibilities of 93.2% and 92% were considered as a good result in order to proceed with the final stage of the experiment - combination of states and with a certain shift between them.

5. Experimental Procedure

The states that constitute the first basis are trivial (Table 1). They can be produced with the help of a single crystal, corresponding to a type I or type II interaction. In the following, we will consider only the generation of the other nine states, i.e. those that form the other three bases. Due to the symmetry of these states, the real parts of the probability amplitudes remain constant and it is only the relative phases between the basic states that are to be varied. This allows us to use the same settings of the waveplates for the generation of nine states. It is also convenient to do three sets of data acquisition - for the fixed φ_{13} values of 0°, $+120^\circ$ and -120°, we change φ_{12} values in the range of a few periods, say from -600 to 600 degrees and measure all tomographic moments for each value of the phase φ_{12} . Then we select the values of φ_{12} that correspond to the generation of the required state. For example, in order to generate state β', we use $\varphi_{13} = -120^\circ$ and $\varphi_{12} = 120^\circ$. The values of the moments at this point allow us to restore a raw density matrix of the generated state and compare it to the theoretical value. According to the Table 2, only four moments (4, 5, 6, 7) corresponding to ImF, ReF, ReD and ImD depend on φ_{12}. It is convenient to perform simultaneous measurements of both real and imaginary parts of the moments F and D in order to ensure that the possible drift of the φ_{12} during measurement results in the same error and can be compensated for afterwards. The values of the other moments (A, B, C, ImE, ReE) remain constant for a given value of φ_{13}. This fact can also be used to perform additional calibration of the setup. For example, the measurement of the relation ImE/ReE and its comparison to theoretical values helps us to verify the correct setting of QP that determine the φ_{13}.

The following procedure can be used in order to verify the orthogonality of a states that form a certain basis. First we choose a fixed state to which we tune our polarization filters. Then we assume that after a beamsplitter each of the photons that form the state goes in to a certain arm of a Brown-Twiss scheme. The values of the angles of waveplates $(\chi_1, \theta_1, \chi_2, \theta_2)$ that assure the maximal projection of the polarization state of each photon on the V direction can be calculated by mapping the fixed state on the Poincaré sphere. We choose states α', α'' and α''' to be our set states for each basis. Then, by setting the phase φ_{13} fixed and by varying the phase φ_{12} we measure the number of coincidence counts that correspond to the certain fourth moment of the field. We will refer to it as

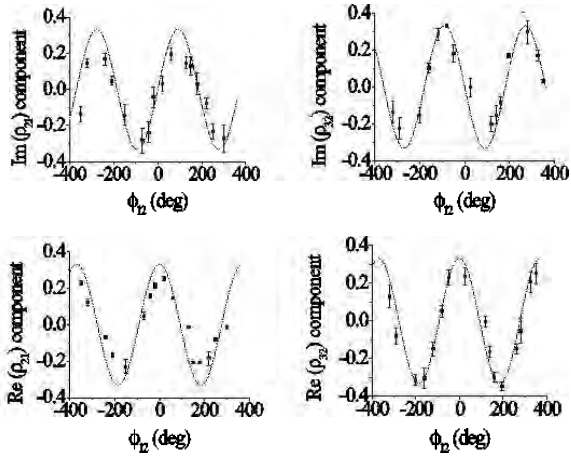


Figure 4.

the "ORT" moment. The dependence of the ORT moment on can be calculated, knowing the set $(\chi_1, \theta_1, \chi_2, \theta_2)$. The coincidence rate falls to zero when the values of φ_{13} and φ_{12} correspond to generation of the states orthogonal to fixed ones.

6. Discussion

Let's consider the generation of the state β''. In this case, $\varphi_{13} = 0°$, $\varphi_{12} = 120°$. On Figure 4 the measured values of the moments are shown, as well as the dependencies of the moments D and F on phase φ_{12} . The phase $\varphi_{13} = 0°$ remained constant during the tomography procedure. After obtaining the dependence of the moments D and F on the phase φ_{12}, we fitted our data with theoretical dependencies, using the least-squares approximation method. The values of the real and imaginary parts of the moments of D and F can be obtained from the fit, as well as the measured values of the other moments that remain constant for $\varphi_{13} = 0°$ were substituted in (5, 6). The reconstructed polarization density matrix ρ_{exp} for state β'' is

$$\rho_{exp} = \begin{pmatrix} 0.35 & -0.05 - 0.21i & 0.32 - 0.01i \\ -0.05 + 0.21i & 0.34 & -0.11 + 0.26i \\ 0.32 + 0.01i & -0.11 - 0.26i & 0.31 \end{pmatrix}.$$

The theoretical density matrix for this state is

$$\rho_{theor} = \begin{pmatrix} 0.33 & -0.17 - 0.29i & 0.33 \\ -0.17 + 0.29i & 0.33 & -0.17 + 0.29i \\ 0.33 & -0.11 - 0.29i & 0.33 \end{pmatrix}.$$

Now we can compute the fidelity F of our experimentally generated state to the theoretical one, using formula $F = \mathrm{tr}\rho_{exp}\rho_{theor}$ and repeat the same proce-

STATE	$\|\alpha'\rangle$	$\|\beta'\rangle$	$\|\gamma'\rangle$	$\|\alpha''\rangle$	$\|\beta''\rangle$	$\|\gamma''\rangle$	$\|\alpha'''\rangle$	$\|\beta'''\rangle$	$\|\gamma'''\rangle$
F_{MLE}	0.9989	0.9967	0.9883	0.9967	0.9983	0.9883	0.9883	0.9989	0, 9967

Table 3.

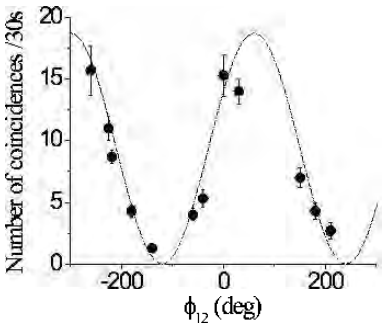


Figure 5.

dure for the rest of the states. For this particular state $F = 0.87$. Note that ρ_{exp} corresponds to the raw experimental data, and can therefore contain some errors that make it non-physical, i.e. its important physical properties such as positivity can be violated. To avoid this, we employ the maximum likelihood method of quantum state root estimation, developed by Bogdanov [36]. The results are presented in Table 3.

The obtained values of fidelities show the high quality of the prepared states. Given the high visibility of the interference patterns that we obtained, we can conclude that our technique for preparation of qutrits is quite reliable. The other practical feature is that all control of the weight of each basic state in superposition (3) is done with the help of linear optical elements, making it easy to switch from one state to another. The other test of the quality of prepared states is the fulfillment of the orthogonality criterion for the states that belong to the same basis. For each fixed state, we calculated the settings of waveplates in our measurement setup that ensure the maximal projection of each photon on a vertical polarization direction. On Figure 5 we show the dependence of a coincidence rate for the given setting of waveplates which correspond to set state $|\alpha''\rangle$. As one can see, for the fixed value of $\varphi_{13} = 0°$ the coincidence rate is almost zero, when phase $\varphi_{12} = -120°$, what corresponds to the generation of the state $|\beta''\rangle$, which is orthogonal to $|\alpha'''\rangle$. The visibility of this pattern is equal to 93.2%. For the other bases, the obtained values of visibilities varied from 92% to 95%. With these values of visibility, the lowest value of coincidence rate corresponds to the accidental (Poissonian) coincidence level and therefore the obtained data verifies the orthogonality criterion. Since we can unambiguously filter the orthogonal state from a fixed one, this suggests that this implementation could be used as a practical realization of a qutrit QKD protocol, based on biphotons. The main cause of the errors was the drift of the phase φ_{12} during the acquisition time. For a given rate of coincidences (1Hz), the average acquisition time for obtaining one point was 30 seconds, and we averaged over three values of a moment. So the measurement time required for the tomography of one state was approximately

1 hour. The phase drift of φ_{12} during that time was about 20 percent from the initial value. Since we measured the real and imaginary parts of moments F and D simultaneously, we can assume that the phase drift was the same for these moments. Therefore it can be compensated for afterwards by comparing our data to the theoretical values of these moments. For a practical implementation of this scheme, such active feedback would be required. The other source of errors that lowered the raw fidelity value was the low coincidence rate and few values taken for averaging. This factor can be eliminated by using two type I crystals instead of type II [35], enabling the use of wider bandwidth filters and larger pinholes, without affecting the visibility of interference (no postselection method).

7. Conclusions

We have discussed an interferometric method of preparing three-level quantum optical systems that relies on the polarization properties of single-mode two-photon light. For adequate state measurement the quantum tomography procedure developed for biphotons-qutrits was applied. A specific sequence of nontrivial states was generated and measured with high fidelity. For some of the states, that form the mutually unbiased bases, we applied the operational orthogonality criterion and experimentally showed the orthogonality of the generated states. This suggests the possible usage of our setup (with some improvements) in a practical realization of a QKD protocol.

This article is based on the experimental works performed in collaboration with M.V.Chekhova, G.A.Maslennikov, L.A.Krivitsky, Yu.I.Bogdanov, M.K.Tey, and A.Penin. The work was supported in part by Russian Foundation for Basic Research (06-02-16769), and National University of Singapore's Eastern Europe Research Scientist and Student Programme.

References

[1] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information, Cambridge Univ. Press (2000).
[2] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Phys. Rev. Lett. **88**, 217901 (2002).
[3] G. M. D'Ariano, in Quantum Optics and Spectroscopy of Solids, edited by A.S.Shumowsky and T.Hakiouglu (Kluwer, Amsterdam), **175** (1997).
[4] S. Schiller, G. Breitenbach, S. F. Pereira, T. Muller, and J. Mlynek, Phys. Rev. Lett., **77**, 2933 (1996).
[5] A. S. Chirkin, A. A. Orlov, D. Yu. Parashuk, Quantum Electronics, **20**, 999 (1993).
[6] V. P. Karassev, J. Sov. Laser Res. **12**, No 5, 147 (1991).
[7] A. V. Masalov, V. P. Karassev. Optics and Spectroscopy, **91**, 558 (2001).
[8] D. James, P. Kwiat, W. Munro, and A. White, Phys. Rev. A, **64**, 052312 (2001).
[9] H.Bechmann-Pasquinucci and A.Peres, Phys.Rev.Lett. **85**, 3313 (2000).
[10] D. Kaszlikowski, D.K.L Oi, M. Christandl, K. Chang, A. Ekert, L.C. Kwek and C.H. Oh, Phys.Rev. A, **67**, 012310 (2003).
[11] D. Bruss, C. Machiavello, Phys. Rev. Lett. **88**, 127901 (2002).
[12] T. Durt, N.J. Cerf, N. Gisin and M. Zukowski, Phys. Rev. A. **67**, 012311 (2003).

[13] D. Collins, N. Gisin, N. Linden, S. Massar, S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).
[14] M. Fitzi, N. Gisin, U. Maurer, quant-ph/0107127.
[15] A. Ambainis, Proc. Symp. Theor. Comput., 01, 134 (2001); A.Ambainis, quant-ph/0204022.
[16] R. T. Thew, A. Acin, H. Zbinden, N. Gisin, Quantum Information and Computation, **4**, No. 2, 093-101 (2004).
[17] A. Vaziri, J.-W. Pan, T. Jennewein, G. Weihs, and A. Zeilinger, Phys. Rev. Lett. **91**, 227902 (2003).
[18] G. Molina-Terriza, A. Vaziri, J. Rehacek, Z. Hradil, A. Zeilinger, Phys. Rev. Lett. **92**, 167903 (2004).
[19] J. C. Howell, A. Lamas-Linares, and D. Bouwmeester, Phys. Rev. Lett. **88**, 30401 (2002).
[20] G. M. D'Ariano, P. Mataloni, M.F. Sacchi, quant-ph/0503227
[21] M. Sullivan-Hale, I. Khan, R. Boyd, J. Howell, Phys. Rev. Lett. **94**, 220501 (2005).
[22] A. V. Burlakov and M.V. Chekhova, Sov. JETP. **75**, 432 (2002).
[23] G. A. Maslennikov, M. V. Chekhova, S. P. Kulik and A. A. Zhukov, Journal of Optics B: Quantum and Semiclassical Optics. **5**, 530 (2003).
[24] A. V. Burlakov, L. A. Krivitsky, S. P. Kulik, et al., Optics and Spectroscopy **94**, 684 (2003).
[25] L. A. Krivitsky, S. P. Kulik, A. N. Penin, and M. V. Chekhova, JETP **97**, 846 (2003).
[26] Yu. Bogdanov, M. Chekhova, S. Kulik, L. Krivitsky, M. K. Tey, C. H. Oh and A. Zhukov. Phys. Rev. A **70**, 042303 (2004).
[27] Yu. I. Bogdanov, L. A. Krivitsky, S. P. Kulik, JETP Lett. **78**, 352 (2003).
[28] D. T. Smithey, M. Beck, M. G. Raymer. Phys. Rev. Lett. **70**, 1244 (1993).
[29] G. M. D'Ariano, M. G. A. Paris. J.Optics B: Quantum Semiclass. Opt. **2**, 113, (2000).
[30] P. Bushuev, V. Karassev, A. Masalov, A. Putilin, Optics and Spectroscopy **91**, 558, (2001).
[31] N. K. Langford, R. B. Dalton, M. D. Harvey, J. L. O'Brien, G. J. Pryde, A. Gilchrist, S. D. Bartlett, A. G. White, Phys. Rev. Lett. **93**, 053601 (2004).
[32] F. De Martini, G. M. D'Ariano, A. Mazzei, and M. Ricci, quant-ph/0207143.
[33] M. V. Chekhova, G. A. Maslennikov and A. A. Zhukov. Sov. JETP Letters. **76**, 696 (2002).
[34] M. V. Chekhova, L. A. Krivitsky, S. P. Kulik, and G. A. Maslennikov, Phys. Rev. A **70**, 053801 (2004).
[35] Y.-H. Kim, M. V. Chekhova, S. P. Kulik, M. H. Rubin, Y. Shih, Phys. Rev. A. 61, 051803(R) (2001).
[36] Yu. Bogdanov, quant-ph/0303014.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Cavity Quantum Electrodynamics: Quantum Entanglement and Information

Jean-Michel Raimond
Laboratoire Kastler Brossel, département de Physique,
Ecole Normale Supérieure,
24 rue Lhomond, F-75005 Paris

Abstract. Cavity Quantum Electrodynamics (CQED) experiments realize the simplest matter-field system: a single atom coupled to a single field mode containing only a few photons. Circular Rydberg atoms and superconducting cavities are ideal tools for CQED, being long-lived and strongly interacting. On the one hand, these experiments can be used to test fundamental features of the quantum world. On the other hand, they implement elementary quantum logic components: qubits and quantum gates. Complex entangled states can be knit by successive atom-field interactions. We review shortly the experiments performed along these lines at ENS. We also discuss perspectives for a quantum information processing architecture based on trapped Rydberg atoms.

Keywords. Cavity quantum electrodynamics, quantum information, quantum entanglement, decoherence

1. Introduction

Quantum information processing puts severe requirements on the physical systems used to implement the elementary information carriers (or qubits). In particular, they should be well isolated from the outside world to avoid the adverse effects of decoherence. At the same time, they must interact strongly together to provide the conditional dynamics of the quantum gates. These two requirements are clearly somewhat contradictory. Few systems, up to now, have been able to fulfill them, at least in part, and to realize quantum gates and elementary quantum logic functions. They are based, in particular, on nuclear spin manipulations by Nuclear Magnetic Resonance [1], superconducting mesoscopic circuits [2], trapped ions [3] or atoms [4] interacting with laser fields, and on Cavity Quantum Electrodynamics [5].

CQED experiments [6] were designed to test intimate features of the quantum world (entanglement, decoherence, measurement theory). They operate on the simplest matter-field system, made up of a single atom interacting with a single field mode containing only a few photons. In the strong coupling regime, the coherent atom-field interaction overwhelms the dissipative processes responsible for decoherence. The CQED experiments thus contain the ingredients of a quantum logic network. The qubits are represented by atomic or field quantum

states. The atom-field interaction generates and processes entanglement between these qubits.

The strong coupling regime can be reached with ordinary atomic transitions interacting with high finesse optical Fabry-Perot cavities [7]. This situation provides a natural interface between flying photonic qubits, which can propagate over long distances in optical fibres, and stationary atomic qubits, which can store and process quantum information. These experiments are well suited, for instance, for the preparation of single photons, which are essential ingredients in quantum key distribution and quantum communication processes [8]. However, they operate on a relatively short time scale, in the microseconds range, making the generation of complex entangled states difficult.

The longest photon storage times, in the millisecond to second range, are obtained in the millimeter-wave domain, with cavities made up of superconducting materials, cooled down to cryogenic temperatures [9]. Circular Rydberg atoms have themselves a long lifetime and an extremely large coupling to millimeter-wave radiation. The combination of these tools provides a nearly ideal situation for CQED. The very long lifetimes of the atomic and field states make it possible to realize complex entanglement manipulation sequences [10].

We shall describe here the millimeter-wave CQED experiments performed at Ecole Normale Supérieure. We first give a short overview of the experimental tools and of the theoretical description of this simple system (Section 2). We discuss quantum information experiments based on the resonant atom-field interaction (Section 3). We then show that a dispersive interaction between the atoms and the cavity can be used to perform a nearly ideal quantum gate operation between two atoms, quite insensitive to cavity damping (Section 4). Finally (Section 5), we shall discuss a promising architecture for quantum information processing based on trapped circular Rydberg atoms.

2. The Circular Rydberg Atoms CQED Set-Up

2.1. Experimental Considerations

Our experimental set-up is sketched in Fig. 1. We only briefly discuss it here. Additional details can be found in [5]. Atoms are prepared in a circular Rydberg state in box B, out of a Rubidium atomic beam effusing from oven O, velocity-selected by laser optical pumping. The circular levels have a high principal quantum number N and maximum orbital and magnetic quantum numbers. In classical terms, the electron is on Bohr's circular orbit, with radius a_0N^2. The quantum orbital is a thin torus centered around this circular orbit.

The atoms can be prepared either in level e (circular state with principal quantum number $N = 51$) or g ($N = 50$). Both levels have a long lifetime (about 30 ms). The $e \rightarrow g$ transition, at $\omega_{eg}/2\pi = 51.1$ GHz, is strongly coupled to the millimeter-wave field due to the large size of Bohr's orbit. The dipole matrix element between these levels is $d = qa_0N^2/2 = 1250$ a.u. The circular atoms are protected from the blackbody background radiation by a 1 K cryogenic environment. The circular states preparation process is time-resolved and the

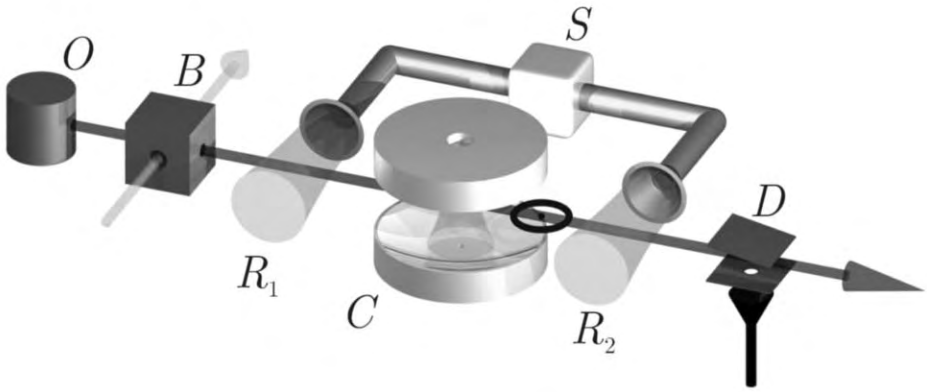


Figure 1. The CQED set-up with circular Rydberg atoms and a superconducting cavity.

atomic samples have a well-defined velocity and initial position. The position of the sample at any time during its further travel in the apparatus is thus well-known, allowing us to apply selective transformations to different samples in the same experimental sequence (individual qubit addressing).

The number of atoms in a sample has a Poisson statistical distribution, with an average value set around 0.1. When an atom is detected, there is a negligible probability for having an unwanted second atom. We thus operate with a reasonable approximation of single atoms at the expense of increased data taking times. When the experimental sequence involves multiple atomic samples, we have to repeat it until a single atom is detected in each sample. The rate of these coincidences decreases exponentially with the number of samples.

The atoms interact with the superconducting cavity C, made up of two spherical superconducting mirrors facing each other. The atoms are nearly resonantly coupled to the TEM_{900} Gaussian mode. The mode waist, $w_0 = 5.96$ mm, is close to the wavelength, $\lambda_{eg} = 6$ mm. The corresponding mode volume, V, [11] is relatively small ($V \simeq 700$ mm^3). The microwave electric field amplitude at cavity center is $E_0 = \sqrt{\hbar\omega_{eg}/2\epsilon_0 V} = 1.5$ mV/m for a single photon stored in the mode. Due to geometrical defects of the mirrors, the degeneracy between the two modes with linear perpendicular polarizations is lifted. The cavity thus sustains two nearly degenerate gaussian modes M_a and M_b, the frequency difference between them being $\delta/2\pi = 128$ kHz. When the atom interacts resonantly with one of these modes, the coupling with the other usually plays a negligible role.

The circular Rydberg atoms experience a second-order Stark shift in an applied electric field. The atomic transition frequency can thus be set, with an excellent time resolution, while the atom crosses the mode. It can be tuned in resonance with either cavity modes, switched rapidly from a resonant interaction to a dispersive one, or even set so far from resonance that the evolution is 'freezed'.

The cavity quality factor Q is an essential parameter. Values as high as 3.10^8, corresponding to a photon lifetime of 1 ms, are obtained by careful polishing and processing of the mirrors. It is limited by diffusion of photons out of the aperture between the mirrors due to the residual roughness of their surface. The photon lifetime is shorter than the atomic lifetime, but still much longer than the time

necessary to create or process atom-cavity entanglement, as shown in the next Section. The strong coupling regime conditions are thus fulfilled.

Some experiments are performed by directly measuring the atomic final energy state with the state-selective field-ionization detector D (detection efficiency greater than 80%). In others, atomic state superpositions are used. These superpositions are prepared, before the atom enters C, by applying to it an auxiliary microwave classical field pulse produced in zone R_1 by the source S (see Fig. 1). Another pulse, in R_2, is used to mix again the atomic energy states after the atom has interacted with C. The pulse R_2 realizes a fully tunable rotation of the Bloch vector representing the atomic transition (a single qubit gate in quantum information terms). The selective detection in D realizes the measurement of the spin along the vertical direction. The combination of this detection with the rotation in R_2 amounts finally to measuring the atomic spin along any axis, providing a full analysis of the final atomic state.

The successive application of the two pulses in R_1 and R_2 constitutes a Ramsey interferometer. In such an atomic interferometer, the probability for finding the atom in a given final state is modulated versus the relative phase of the two pulses, which can easily be controlled. We study how the fringe phase and amplitude are affected by the presence of photons in the cavity and we gain in this way useful information on the atom-photon interaction process.

2.2. The Spin/Spring System

The atom-cavity system can be described with very simple theoretical tools. The two-level atom is a realization of the spin-1/2 system. The single cavity mode is analogous to a single dimensional harmonic oscillator. This spin/spring system is ubiquitous in physics. It describes as well, for instance, the dynamics of a trapped ion. In this case, the spin is made up of the two internal levels of the ion, the spring is the harmonic motion in the trap and the coupling is provided by tailored laser beams impinging onto the ion [12].

The complete hamiltonian of the atom-cavity system is:

$$H = H_a + H_c + H_{ac} \ , \tag{1}$$

where $H_a = \hbar\omega_{eg}\sigma_Z/2$ and $H_c = \hbar\omega_c a^\dagger a$ are the atom and cavity hamiltonians. The cavity mode angular frequency is ω_c. We get rid here of the vacuum energy $\hbar\omega_c/2$. The atomic operator σ_Z is the Pauli matrix associated with the two-level transition $|e\rangle \rightarrow |g\rangle$, a and $a^\dagger$ are the photon annihilation and creation operators, checking the bosonic commutation rule $\left[a, a^\dagger\right] = 1$.

The atom-field coupling hamiltonian is, with appropriate phase conventions:

$$H_{ac} = -i\hbar\frac{\Omega_0}{2}\left[a\sigma_+ - a^\dagger\sigma_-\right] \ , \tag{2}$$

where we introduce the 'single photon Rabi frequency', Ω_0, defined by:

$$\Omega_0 = 2\frac{d\mathcal{E}_0}{\hbar} \ . \tag{3}$$

With the experimental parameters given above, $\Omega_0/2\pi = 50$ kHz. The coupling H_{ac} describes processes in which the atom and the cavity exchange a quantum of excitation. The angular frequency Ω_0 of this exchange measures the strength of the atom-field coupling. It is proportional to the interaction energy of the atomic dipole with a classical field corresponding to a single photon stored in the cavity.

The 'uncoupled states', eigenstates of $H_a + H_c$, are the tensor products $|e, n\rangle$ and $|g, n\rangle$ of atomic energy states and cavity Fock states. At resonance ($\omega_c = \omega_{eg}$), the uncoupled states $|e, n\rangle$ and $|g, n+1\rangle$ are degenerate and coupled by H_{ac}. Note that the ground state of the system, $|g, 0\rangle$, is non-degenerate and not coupled to any other state. It is thus impervious to atom-cavity interaction. The complete hamiltonian breaks down in two-dimensional matrices in the $\{|e, n\rangle, |g, n+1\rangle\}$ subspaces and can be explicitely diagonalized, providing the 'dressed' atom-field levels [11]. For quantum information purposes, we can restrict our study to the two states with a single excitation, $|e, 0\rangle$ and $|g, 1\rangle$. At resonance, the dressed states $|\pm, 0\rangle$ have an energy separation $\hbar\Omega_0$ and are symmetric and antisymmetric superpositions of the uncoupled states:

$$|\pm, 0\rangle = \frac{1}{\sqrt{2}} \left[|e, 0\rangle \pm i|g, 1\rangle\right] \ . \tag{4}$$

2.3. Quantum Rabi Oscillation

Let us consider the simple case of an atom entering the empty cavity in state $|e\rangle$. The initial joint state is $|e, 0\rangle$, which can be written as the sum of the two dressed levels. These non-degenerate energy eigenstates evolve in time with different Bohr frequencies. The final state reached after an interaction time t is then:

$$|\Psi_e(t)\rangle = \cos\frac{\Omega_0 t}{2}|e, 0\rangle + \sin\frac{\Omega_0 t}{2}|g, 1\rangle \ . \tag{5}$$

It oscillates between $|e, 0\rangle$ and $|g, 1\rangle$ at the vacuum Rabi frequency Ω_0. When the atom is initially in state $|g\rangle$ in a cavity containing 1 photon, the complete state at time t is instead:

$$|\Psi_g(t)\rangle = -\sin\frac{\Omega_0 t}{2}|e, 0\rangle + \cos\frac{\Omega_0 t}{2}|g, 1\rangle \ . \tag{6}$$

This quantum Rabi oscillation can be viewed as an 'oscillatory spontaneous emission' process. An atom initially in $|e\rangle$ emits a photon while undergoing a transition to $|g\rangle$. In free space, the photon escapes at light velocity and the atom remains in state $|g\rangle$. This is the usual irreversible spontaneous emission. In the cavity, the photon emitted by the atom remains trapped for a long time. It can be re-absorbed, then emitted again. In more mathematical terms, the coupling to a continuum of modes (free space case) is replaced here by the coupling to a single quantum oscillator. The irreversible damping is thus replaced by a coherent oscillation. The observation of this coherent atom-field coupling requires atomic and cavity field lifetimes much longer than the Rabi period ('strong coupling regime' of CQED), a condition which is achieved with circular Rydberg atoms and superconducting cavities.

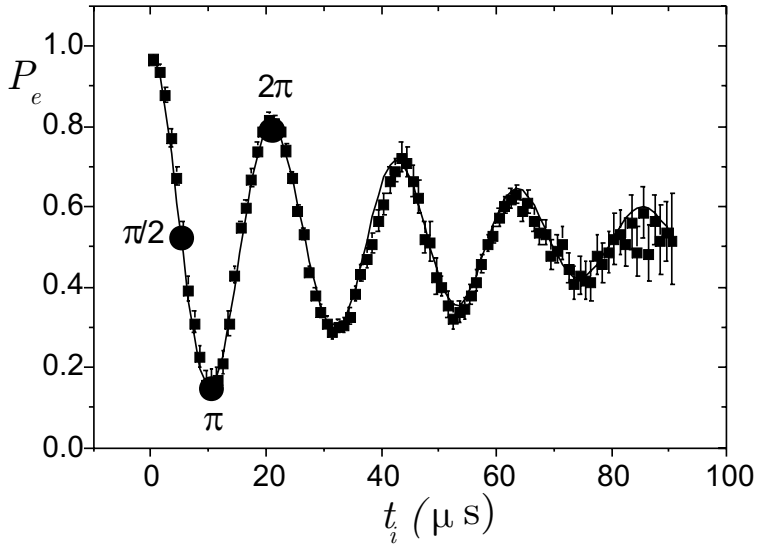


Figure 2. Rabi oscillation signal: A single atom emits and reabsorbs a single photon. The probability for detecting the atom in $|e\rangle$ is plotted as a function of the atom-cavity interaction time. Interaction times corresponding to $\pi/2$, π and 2π pulses used in quantum information processing experiments are marked by solid black circles.

Figure 2 presents the experimental Rabi oscillation signal for an atom entering the empty cavity in $|e\rangle$. The atom-cavity interaction time t is tuned by selecting the atomic velocity and making use of the Stark atomic frequency shift. Four complete oscillations are visible, a clear indication of the strong coupling regime. More importantly for our purpose, peculiar interaction times marked in figure 2 provide us with the basic stitches required to knit complex entangled states.

3. Resonant Atom-Cavity Entanglement Manipulations

3.1. Quantum Stitches

Let us start by considering the '$\pi/2$-quantum Rabi pulse' corresponding to $\Omega_0 t = \pi/2$. The atom-cavity interaction then realizes the transformation:

$$|e,0\rangle \longrightarrow \Psi_e(\pi/2\Omega_0) = \frac{1}{\sqrt{2}}\left[|e,0\rangle + |g,1\rangle\right] \ . \tag{7}$$

The atom and the cavity form now an entangled pair, reminiscent of the EPR situation [13]. The typical interaction time needed to produce this entanglement is a quarter of the quantum Rabi oscillation period, i.e. 5 μs. This is much shorter than the lifetime of the entangled state, 1 ms, limited mainly by the photon decay. The atom-field entanglement thus survives the separation of these two systems and exhibits all the weird features of non-local quantum states.

An effective interaction time such that $\Omega_0 t = \pi$ defines a π-quantum Rabi pulse corresponding to the transformations:

$$|e,0\rangle \longrightarrow |g,1\rangle \quad \text{and} \quad |g,1\rangle \longrightarrow -|e,0\rangle \ . \tag{8}$$

The excitation is exchanged between the atom and the cavity. More generally, when the atom enters the empty cavity in an arbitrary superposition of $|e\rangle$ and $|g\rangle$, $c_e|e\rangle + c_g|g\rangle$, the state transformation reads:

$$(c_e|e\rangle + c_g|g\rangle)\,|0\rangle \longrightarrow |g\rangle\,(c_e|1\rangle + c_g|0\rangle) \ . \tag{9}$$

The atom always ends up in level $|g\rangle$ and the cavity is prepared in a quantum superposition of $|0\rangle$ and $|1\rangle$. In quantum information terms, the qubit initially carried by the atom is deposited into the cavity [14]. This information exchange is reversible. Sending later in the cavity an atom in state $|g\rangle$, the same π-pulse interaction time leads to the transformation:

$$|g\rangle\,(c_e|1\rangle + c_g|0\rangle) \longrightarrow (-c_e|e\rangle + c_g|g\rangle)\,|0\rangle \ . \tag{10}$$

The π-quantum Rabi pulse is an essential ingredient in these experiments. Only the atom is directly observable by the field-ionization detector D. The field remains hidden in the cavity. The information we get on it must be carried away by a probe atom. The field state is copied onto the probe by π-quantum Rabi pulse, as described above.

Finally, after a 2π-quantum Rabi pulse ($\Omega_0 t = 2\pi$), the system returns to the initial level, with a global phase shift of the atom-field state:

$$|e,0\rangle \longrightarrow -|e,0\rangle \quad \text{and} \quad |g,1\rangle \longrightarrow -|g,1\rangle \ . \tag{11}$$

These transformations are reminiscent of the π-phase shift of a spin-1/2 undergoing a 2π rotation in real space. The atom-cavity phase shift is conditioned to the atom and cavity states. It occurs in the state $|g,1\rangle$, while $|g,0\rangle$ remains invariant. This conditional property can be used to realize a quantum logic gate with atoms and cavities [15].

We have used these three basic quantum stitches to process quantum information in an increasingly complex way: from a quantum memory [14], storing for a while quantum information in the cavity, to the step-by-step generation of a three-particle entangled state [10], through the generation of a two-field mode entangled state [16] or of an atomic EPR pair [17]. We now briefly describe some of these experiments.

3.2. A Quantum Phase Gate

The quantum phase gate (QPG) transformation simply reads:

$$|a,b\rangle \longrightarrow \exp(i\phi\delta_{a,1}\delta_{b,1})|a,b\rangle \ , \tag{12}$$

where $|a\rangle$, $|b\rangle$ stand for the basis states ($|0\rangle$ or $|1\rangle$) of the two qubits and $\delta_{a,1}$, $\delta_{b,1}$ are the usual Kronecker symbols. The QPG leaves the initial state unchanged, except if both qubits are 1, in which case the state is phase–shifted by an angle ϕ.

In order to implement the QPG, let us now consider a third atomic level $|i\rangle$, the circular Rydberg state with principal quantum number 49. Due to the large

detuning of the $|i\rangle \rightarrow |g\rangle$ transition with the cavity mode frequency, this level is not coupled to C. The transformation corresponding to the 2π Rabi pulse in C is:

$$\begin{array}{l} |i,0\rangle \longrightarrow |i,0\rangle \\ |i,1\rangle \longrightarrow |i,1\rangle \\ |g,0\rangle \longrightarrow |g,0\rangle \\ |g,1\rangle \longrightarrow -|g,1\rangle \end{array} \tag{13}$$

This is exactly the $\phi = \pi$ QPG, when the states $|i\rangle$ and $|g\rangle$ are mapped onto the logical $|0\rangle$ and $|1\rangle$ states of the atomic qubit. The ability of this gate to generate entangled states can be demonstrated by operating it on a superposition state of either the atomic or field qubit [15]. As an example, from the initial atom-field state $1/2(|i\rangle + |g\rangle)(|0\rangle + |1\rangle)$, the QPG prepares the maximally entangled state:

$$\frac{1}{2}(|i\rangle + |g\rangle)(|0\rangle + |1\rangle) \longrightarrow \frac{1}{2}(|i\rangle + |g\rangle)|0\rangle + (|i\rangle - |g\rangle)|1\rangle) \ . \tag{14}$$

After interaction with C, the atomic state superposition is phase-shifted by π if and only if the cavity contains one photon. Note that the 2π pulse interaction with C leaves the photon number unchanged: when the atom is in $|g\rangle$ it absorbs the photon and reemits it.

Measuring the phase of the atomic superposition state thus amounts to a Quantum Non Demolition (QND) detection of a single photon in C. The photon is still in the cavity after its detection, a quite unusual situation. As shown in [18], this measurement can be implemented using the Ramsey interferometer between zones R_1 and R_2. This experiment demonstrates that the phase of an atomic superposition is coherently controlled by the state of a single photon. Symmetrically, we have checked that the phase of a superposition of the $|0\rangle$ and $|1\rangle$ field states is shifted by π under the operation of the QPG when the atom is prepared in $|g\rangle$ [15]. These two experiments check the coherent operation described by Eq.(14).

Note that the QND photon measurement procedure leads also to a direct determination of the Wigner distribution at the origin of phase space for the vacuum state or for a one-photon field state [19]. A complete determination of the cavity-field Wigner function over the whole phase space has recently been realized [20]. It extends the QND measurement principle to arbitrary photon numbers by using the dispersive atom-field interaction instead of the resonant one [21]. This Wigner function measurement procedure provides an extremely useful tool for analyzing in depth the quantum state of the cavity qubit.

3.3. Multiparticle Entanglement

The most complex quantum operation realized so far is the generation of a three-qubit entangled state. We send across C, initially empty, an atom A_1 initially in $|e\rangle$. A $\pi/2$-quantum Rabi pulse prepares the state $(|e,0\rangle + |g,1\rangle/\sqrt{2}$. We then send a second atom A_2. Initially in $|g\rangle$, it is prepared, before C, in the state $(|g\rangle + |i\rangle)/\sqrt{2}$ by a classical pulse. This atom interacts with C for a 2π-quantum Rabi pulse and performs the QPG operation. The resulting $A_1 - A_2 - C$ quantum state is :

$$|\Psi_{\rm triplet}\rangle = \frac{1}{2}\left[|e_1\rangle(|i_2\rangle + |g_2\rangle)|0\rangle + |g_1\rangle(|i_2\rangle - |g_2\rangle)|1\rangle\right] \tag{15}$$

(the indices correspond to the atom number). Eq.(15) describes a three-particle entangled state and can be rewritten as :

$$|\Psi_{\rm triplet}\rangle = \frac{1}{2}\left[|i_2\rangle(|e_1,0\rangle + |g_1,1\rangle) + |g_2\rangle(|e_1,0\rangle - |g_1,1\rangle)\right] \ , \tag{16}$$

describing an $A_1 - C$ EPR pair whose phase is conditioned to the state of A_2.

Since $|\Psi_{\rm triplet}\rangle$ involves two levels for each subsystem, it is equivalent to an entangled state of three spins 1/2. Let us define the states $|+_i\rangle$ ($|-_i\rangle$) (with $i = 1, 2$) as $|+_1\rangle = |e_1\rangle$ ($|-_1\rangle = |g_1\rangle$), $|\pm_2\rangle = (|g_2\rangle \pm |i_2\rangle)/\sqrt{2}$ and $|+_C\rangle = |0\rangle$ ($|-_C\rangle = |1\rangle$). With these notations, $|\Psi_{\rm triplet}\rangle$ takes the form of the Greenberger, Horne and Zeilinger (GHZ) spin triplet [22]:

$$|\Psi_{\rm triplet}\rangle = \frac{1}{\sqrt{2}}\left(|+_1,+_2,+_C\rangle - |-_1,-_2,-_C\rangle\right) \ . \tag{17}$$

The entanglement is checked by performing measurements of the three qubits in orthogonal bases. First, the state of the cavity is copied onto a third atom, A_3, initially in $|g\rangle$ and interacting with the cavity for a π-quantum Rabi pulse. Classical resonant pulses on the atoms make it possible, as discussed above, to detect the associated spin-1/2 in any direction.

The analysis of the experimental results shows that the three-particle state preparation fidelity is about 54%, well in the quantum domain [10]. One of the main limitations to the fidelity is due to the cavity losses.

Let us mention that the production and characterization of this GHZ state is a quite complex quantum logic process. It involves four qubits (three atoms and the cavity mode), undergoing three single-qubit gates (resonant microwave pulses) and three two-qubit gates (atom-cavity interaction). It is still among the most complex quantum information sequences realized so far on individually addressed isolated qubits.

4. Direct Atom-Atom Entanglement: Cavity-Assisted Collisions

The atom-atom entangling procedures outlined above rely on the exchange of a photon between the atom and the cavity. The quantum information is transiently stored as a superposition of the zero- and one- photon states. These schemes are thus sensitive to cavity losses, the main cause of decoherence in our experiments (the atomic lifetime being much longer than the cavity damping time).

It is possible to circumvent this problem by entangling two atoms directly, in a collision process assisted by the non-resonant cavity modes [23,24]. The first atom (A_1) is initially in $|e\rangle$ and the second (A_2) in $|g\rangle$. The atoms have different velocities, so that the second catches up the first at cavity center, before exiting first from C. The two cavity modes M_a and M_b are detuned from the $|e\rangle \rightarrow |g\rangle$ transition frequency by amounts Δ and $\Delta + \delta$, greater than Ω_0. Due to energy conservation, real photon emission cannot occur in this case. Atom A_1 can, however, virtually emit a photon immediately reabsorbed by A_2. This leads to a Rabi

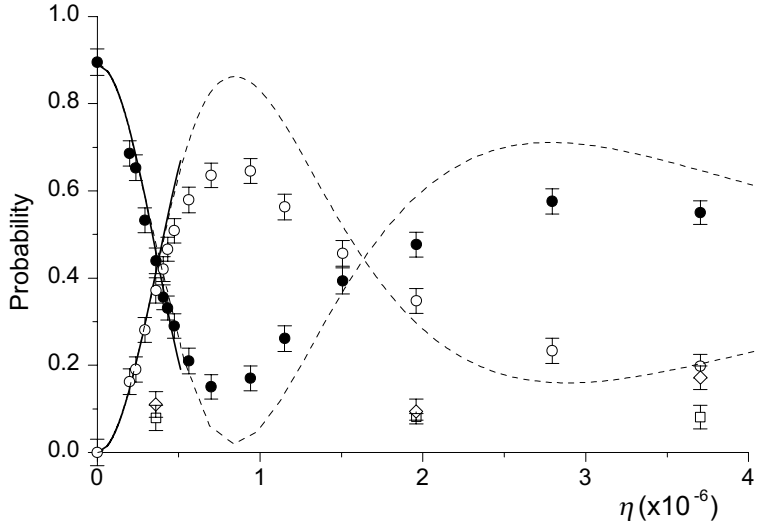


Figure 3. Cavity assisted collision. Experimental values of the joint detection probabilities P_{eg} (solid circles) and P_{ge} (open circles) versus the parameter η. The open squares and triangles correspond to the spurious channels for which both atoms are detected in the same state. Their population remains below 10%. Solid lines for small η values correspond to a simple analytical model based on second order perturbation theory. The dashed lines (large η) present the results of a numerical integration of the system evolution.

oscillation between states $|e,g\rangle$ and $|g,e\rangle$ and thus to atom/atom entanglement generation for most interaction times.

The situation is reminiscent of a resonant van der Waals collision in free space, which can also produce atom-atom entanglement for small enough impact parameters [25]. In the present case, the detuned cavity modes considerably enhance the atom-atom interaction. Note that, in this peculiar 'collision' process, the actual distance between the atoms is irrelevant, provided they both interact with the modes.

The quantum amplitudes associated to states $|e,g\rangle$ and $|g,e\rangle$ are periodic functions of the collision duration (which depends on the atomic velocities). The oscillation frequency associated to this second-order collision process is $(\Omega_0^2/4)[1/\Delta + 1/(\Delta+\delta)]$. By repeating the experiment, we reconstruct the probabilities P_{eg} and P_{ge} for finding finally the atom pair in states $|e,g\rangle$ and $|g,e\rangle$. We plot these probabilities versus the dimensionless parameter $\eta = \omega_{eg}[1/\Delta + 1/(\Delta+\delta)]$ (see Fig. 3). The oscillations of P_{eg} and P_{ge} as a function of η are well accounted for by theoretical models (solid and dashed lines in Fig. 3).

We have realized the situation of maximum entanglement by adjusting η to the lowest value corresponding to $P_{eg} = P_{ge} = 0.5$. The collision should then result in the preparation of an atomic EPR pair. We have checked the coherent nature of the pair by performing again measurements of the two atomic qubits in non-compatible bases.

Since this entanglement procedure implies only a virtual photon exchange with the detuned cavity mode, it is, in first order, insensitive to the cavity damping or to a stray thermal field in the modes. It thus opens interesting perspectives

for demonstrating elementary steps of quantum logic with moderate Q cavities at finite temperature.

We have shown theoretically that the two-qubit Grover search algorithm [26] could be realistically implemented in our set-up with two cavity-assisted collisions between two atoms, performed during the common interaction of the atoms with the cavity mode [27].

5. An Architecture for Quantum Information Processing

The cavity-enhanced collision realizes gate operations with a moderate cavity quality and an unlimited fidelity (at least in principle). However, a large fidelity implies a large atom-cavity detuning and hence a long atom-field interaction time. Such long times can only be obtained through a complete control over the atomic motional degrees of freedom. Another limitation of the experiments described above is the random circular states excitation process. We mimic single atom samples with low average value Poisson distributions, resulting in data taking times which grow exponentially with the number of samples involved in the experimental sequence.

Complex ('scalable') entanglement manipulations thus require a deterministic preparation of cold and trapped circular atoms. The trap must increase the atomic lifetime beyond the 30 ms spontaneous emission limit. It must also be compatible with the preservation of a qubit coherence, coded onto a superposition of two circular states, $|e\rangle$ and $|g\rangle$ for instance. We have recently shown that a coherence-preserving trap could be designed, providing a scalable architecture for quantum computing [28]. We think it interesting to describe it shortly, since it relies on an interesting combination of cavity QED effects.

First, we get rid of spontaneous emission. The only relaxation channel for the circular state $|e\rangle$ is a transition towards $|g\rangle$ by emission of a photon at the millimeter wavelength λ_{eg}. The level $|g\rangle$, in turn, decays towards $|i\rangle$, emitting a photon at $\lambda_{gi} < \lambda_{eg}$. Both these decay processes can be blocked. The circular atom is placed between two parallel conducting planes, separated by a distance d such that $d < \lambda_{gi}/2 < \lambda_{eg}/2$. A d.c. voltage applied across the mirrors creates an electric field parallel to the normal of the mirrors (z axis). The atomic dipole is then perpendicular to z and the emitted microwave photon circularly polarized with respect to z. This polarization is not compatible with the mirrors, which are a waveguide below cut-off for both atomic transition frequencies. There is no mode of the radiation field for the atom to emit into. Ideally, the atom stays in $|e\rangle$ or $|g\rangle$ for an infinite time!

In a real situation, two imperfections contribute to a residual emission rate. The mirrors are not perfect. The electric field experienced by the atom may also be tilted with respect to the z axis, by a small angle θ (an important effect in the trap described below). The residual spontaneous emission rate is then $\Gamma_c = \Gamma_a + \Gamma_s$, where Γ_s is the residual rate due to mirrors imperfections and $\Gamma_a = \Gamma \sin^2 \theta$ ($\Gamma = 33\ \mathrm{s}^{-1}$ is the free-space spontaneous emission rate).

In the millimeter-wave range, lifetime enhancements by an order of magnitude have been observed in the early times of CQED [29,30] for circular Rydberg

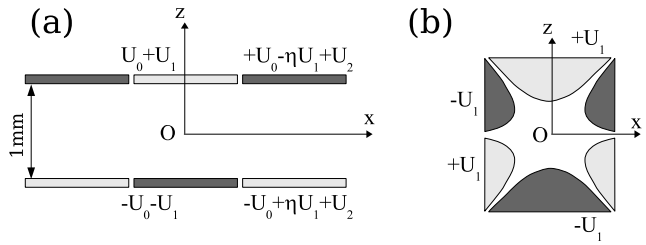


Figure 4. Electric trap geometries. Sections in the xOz plane, with applied potentials. The electrodes are shaded according to the phase of the oscillating potential U_1. Both traps have a cylindrical symmetry around Oz (a) Trap compatible with spontaneous emission inhibition. The plates spacing and the diameter of the inner electrodes are 1 mm. The a.c. potential applied on the outer electrodes is ηU_1. The η factor, dependent on the electrodes geometry, is tuned to optimize the trap. (b) Electrode geometry creating an exact hexapolar potential.

atoms or for the quantized cyclotron motion of a trapped electron in a magnetic field. Lifetimes in the few tens of seconds range, corresponding to a 10^3 enhancement, seem possible with high quality mirrors. The tilt angle θ should thus remain below 10 mrad. The temperature should also be kept in the Kelvin range. The circular atoms, in $|g\rangle$ for instance, may absorb thermal photons with a π polarization (compatible with the mirrors geometry), in a transition towards a non-circular level in the $N = 51$ manifold. These processes are efficiently suppressed at cryogenic temperatures.

Spontaneous emission inhibition does not prevent us from directly addressing the $|e\rangle \rightarrow |g\rangle$ transition for the realization of single-qubit gates. Evanescent field modes in the waveguide below cut-off can be excited by a strong classical source through a small iris pierced in one of the mirrors. Note also that the atom/mirrors distance is of the order of a millimeter. The van der Waals resonant interaction of the atom with its electric images in the mirrors corresponds then to small atomic frequency shifts, in the Hertz range.

The Rydberg trap must achieve a strong confinement near the centre of the waveguide and preserve a tight alignment of the directing field on the z axis. We propose to use an electric field trap, based on the quadratic Stark polarizability of circular levels, inspired by a recent proposal for a ground state atom trap [31]. Levels $|e\rangle$ and $|g\rangle$ are shifted by about -2 MHz/(V/cm)^2 in an electric field. They are attracted towards high fields. Since there is no maximum of an electrostatic field in free space, there is no stable static electric trap. The same problem is encountered for charged particles, ions or electrons. It is solved by adding of a magnetic field (Penning trap) or by using a dynamic trap in an alternating quadrupolar potential (Paul trap). Only the latter solution is appropriate the present context.

A quadrupolar Stark energy derives from an hexapolar electrostatic potential. The simplest electrode configuration producing such a potential is represented in figure 4(b). It is clearly not appropriate for spontaneous emission inhibition. We thus 'flatten' the structure and use instead the geometry presented in figure 4(a).

The trap is made up of two parallel planes, with circular electrodes cut out in their centers. The potentials applied to the electrodes are shown in figure 4(a). The static potential U_0 creates a large, homogeneous directing electric field

aligned with the Oz axis. The a.c. potential U_1 creates, near the trap centre O, an approximately hexapolar potential oscillating at the frequency ω_1, in the kiloHertz range. The potential U_2, finally, creates a quadrupole potential and a nearly constant vertical force near trap center, adjusted to compensate for gravity (assumed to be antiparallel with Oz).

The trap characteristics can be determined by numerical integration of the atomic motion. For instance, for $U_0 = 0.2$ V, $U_2 = -0.003$ V, $U_1 = 0.155$ V and $\omega_1/2\pi = 430$ Hz, we get a stable trapping. The atomic motion is the superposition of a 'micro motion' at the trap drive frequency and of a nearly harmonic 'macro motion', with oscillation frequencies of 175 Hz and 64 Hz along Oz and Ox. They correspond to a trap depth in the 100 μK range, well adapted to standard laser cooling procedures. The atoms explore a 100 μm region around the trap centre. In this region, the electric field is predominantly the homogeneous component produced by U_0. The maximum angle between the local time-varying electric field and the Oz axis is $\theta = 10$ mrad, corresponding to a residual emission rate $\Gamma_a = 10^{-4}\Gamma$, i.e. to a 30 s lifetime. Trapping atoms for time intervals in the minute range is thus feasible.

This trap is not yet compatible with the preservation of a qubit coherence, coded on a superposition of $|e\rangle$ and $|g\rangle$. These levels have different Stark polarizabilities [the transition frequency shift is -255 kHz/(V/cm)^2]. Each atom has a different trajectory in the trap and experiences a unique electric field history. This results in an inhomogeneous broadening of the transition and to a rapid blurring of the coherence phase. Taking as initial condition atoms with a 0.3 μK temperature at trap centre, we find that atomic coherence is lost after ten microseconds only.

This coherence loss effect can be considerably reduced by making the Stark polarizability state-independent, using a microwave dressing of level $|g\rangle$, as illustrated in figure 5(a). A microwave field polarized parallel to the quantization axis z couples $|g\rangle$ to a level $|e'\rangle$, in the same hydrogenic manifold as $|e\rangle$. The polarization of this dressing field is compatible with the limiting conditions imposed by the trap electrodes. Level $|e'\rangle$ has a linear Stark effect and thus a much higher polarizability than $|e\rangle$ or $|g\rangle$. The dressing field is non-resonant (detuning $\delta_{ge'}$ with respect to the $|g\rangle \rightarrow |e'\rangle$ transition in the average trapping electric field E_a). It does not induce any transition between $|g\rangle$ and $|e'\rangle$, but admixes a little bit of the high polarizability of $|e'\rangle$ into $|g\rangle$. By choosing properly the detuning $\delta_{ge'}$ and the Rabi frequency $\Omega_{ge'}$, it is possible to equalize the slopes and nearly equalize the curvatures of the Stark energy of $|e\rangle$ and $|g\rangle$ in the explored electric field range. The variation of the $|e\rangle \rightarrow |g\rangle$ transition frequency over the motion in the trap is now only a few Hertz, as shown in figure 5(b), making it possible to preserve a qubit coherence over times in the tens of milliseconds range.

We have numerically followed the phase of an $|e\rangle/|g\rangle$ coherence over a set of typical trajectories in the trap. One of these evolutions is presented in figure 5(c). The phase evolves nearly linearly with small steps around the times when the atom is at an outer turning point of its trajectory, experiencing then a maximum field amplitude. The step size, $\Delta\phi$, remains nearly constant over seconds.

Such inhomogeneous phase spreadings can be efficiently combated by an echo method. We simulate a Ramsey experiment. At time $t = 0$, a $\pi/2$ classical pulse

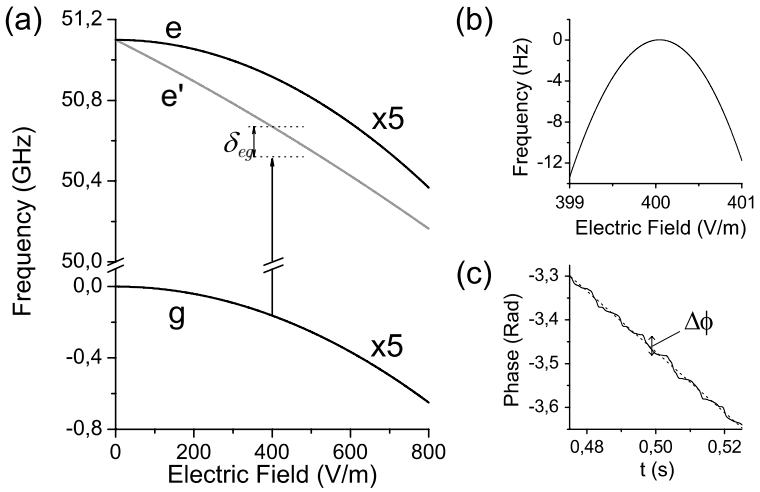


Figure 5. (a) Stark energy of levels $|e\rangle$, $|g\rangle$ and $|e'\rangle$ versus electric field. States $|e\rangle$ and $|g\rangle$ have, in this field range, an approximately quadratic Stark effect, magnified here by a factor of 5 for the sake of clarity. Level $|e'\rangle$ has a much larger, linear Stark effect. The microwave dressing (vertical arrow) predominantly mixes $|g\rangle$ and $|e'\rangle$. (b) variation of the $|e\rangle \rightarrow |g\rangle$ transition frequency around the average electric field E_a in presence of an optimal dressing ($\Omega_{ge'}/2\pi = 200$ MHz, $\delta_{ge'}/2\pi = 556.23$ MHz). (c) Solid line: phase accumulated by an $|e\rangle/|g\rangle$ coherence over a 40 ms time period in the presence of the microwave dressing showing a step structure (average step size $\Delta\phi$), in phase with the atomic motion. Dashed line: linear fit (slope 6.9 rad/s) over a total time of 1s.

resonant on the $|e\rangle \rightarrow |g\rangle$ transition creates an atomic coherence. At a time $T_\pi = 0.5$ s, we apply a π resonant Rabi pulse which exchanges the levels. This amounts to changing the sign of the accumulated phase of the e/g coherence. From this time on, the coherence's phase continues to drift with the same linear slope as before. The phase thus returns to zero, a value reached within a step size $\Delta\phi$ at time $2T_\pi$. The cancelation of the accumulated phase occurs for all atomic trajectories, and the coherence is restored. We apply a second $\pi/2$ pulse at time t that probes this coherence.

The contrast of the calculated Ramsey fringes, $\mathcal{C}$, is presented versus t in figure 6. The dotted line corresponds to the situation without the echo pulse at time T_π. The contrast irreversibly decays, with a short time constant. The solid line corresponds to the signal with echo. The contrast revives sharply around $2T_\pi$. The peak value, around 82%, corresponds to a five seconds lifetime for the atomic coherence, a quite remarkable result.

Finally, the 'dipole blockade' mechanism [25,32] makes it possible to prepare deterministically a single Rydberg atom in a low angular momentum state. This atom can then be efficiently transferred to the circular state and caught in the electric trap. The starting point of the excitation process is a dense cloud, a few microns in size, containing a few hundred ground state atoms, held at a low temperature in a tightly confining magnetic trap. A laser, resonant with the transition between the ground state and a Rydberg level, can then easily excite a first atom. The strong dipole-dipole interaction between the Rydberg atoms strongly shifts the transition frequency towards a state with two Rydberg atoms.

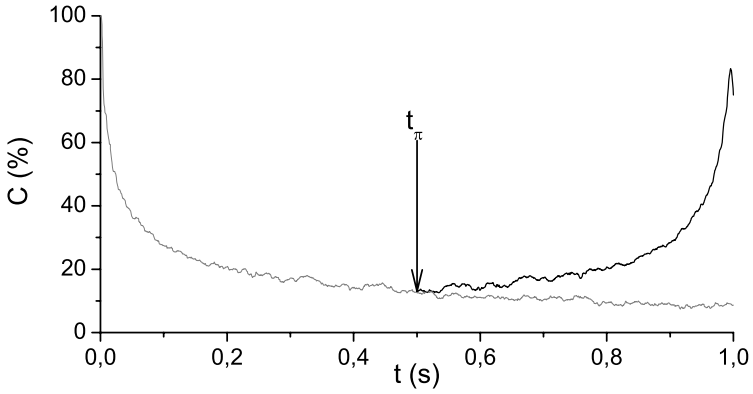


Figure 6. Simulated Ramsey fringes contrast $\mathcal{C}$ as a function of the time interval t between the two $\pi/2$ Ramsey pulses. Dotted line: contrast decay without echo, averaged over 10 000 trajectories. Black line: contrast for the same trajectories with a π pulse at $T_\pi = 0.5$ s. The sharp contrast revival at $t_f = 1$ s peaks at 82%.

A single atom can thus be prepared. With a proper tailoring of the laser excitation, the excitation efficiency can be high, leading to the deterministic preparation of a single circular atom.

The initial atom cloud must be cold and dense. The required tight confinement is typical of the 'atom chips' experiments, in which the trapping magnetic field is provided by a set of micron-sized conductors produced by lithographic techniques on an insulating chip [33]. Since we need low temperatures for the sake of Rydberg states stability, it is tempting to use a superconducting material for the realization of the atom chip. Power dissipation, a major issue for room-temperature chips using normal metals, is then suppressed. Permanent supercurrents in loops also provide the high magnetic field stability required for an efficient trapping and a coherent transport of atomic samples, opening interesting perspectives for an on-chip atomic interferometer.

We thus have in mind a completely integrated superconducting device. Two chips facing each other integrate the magnetic trap for ground state atoms, electrodes and conductors providing the electric and magnetic fields needed for the circular states preparation process and for the electric trap. They may even incorporate electron detectors for the final read-out of the atomic state [34].

Atom-atom entanglement in this device can be produced by the direct van der Waals interaction between two atoms held in neighbouring traps. It could also result from a cavity-assisted collision. Macroscopic Fabry Perot cavities are not compatible with the tight geometry of the trap. On-chip planar transmission line cavities are much more appealing. They have recently been used in a remarkable experiment demonstrating the strong coupling regime of CQED with artificial atoms, made up of tiny superconducting circuits [35]. These cavities provide a high field confinement. Their quality factor is moderate (10^4 to 10^6), but large enough for an efficient assisted collision.

6. Conclusion and Perspectives

The circular Rydberg atoms already made it possible to operate interesting quantum entanglement processing sequences [5]. The extension to much more complex algorithms requires improvements of the present set-up. Two major extensions are now in progress. On the one hand, the atom-cavity experiment will be extended to incorporate two high-Q cavities. On the other hand, we are setting up the Rydberg atom-chip experiment.

For the cavity QED set-up, the fidelity it limited by the imperfections of the elementary gates and by the cavity losses. A better control of the stray fields in the set-up, which seem to be a major cause of imperfections, could improve noticeably this fidelity. The cavity losses could be reduced with a new mirror technology. We have recently obtained a 14 ms damping time for a cavity, which opens very interesting perspectives for the tailoring of complex entangled state.

The set-up under construction will incorporate two cavities. This offers an additional flexibility for quantum information processing. The two cavities could be used, for instance, as teleportation booths transmitting an atomic state at a distance [36]. The first cavity could also entangle a few atoms, coding a quantum error correction code, that the second cavity would read out.

Let us note also that these experiments are well-suited for the exploration of other basic quantum mechanisms essential for quantum information processing. In particular, mesoscopic coherent fields stored in the cavity provide an unprecedented tool for an in-depth study of the decoherence mechanisms [37,38]. The two-cavity set-up makes it possible to generate and study non-local mesoscopic states (EPR pairs made of mesoscopic cavity fields) and allows new tests of our understanding of the decoherence process.

The Rydberg chip experiment opens fascinating perspectives for quantum entanglement manipulation. It circumvents many of the limitations of the set-up using randomly prepared 'flying' atoms and will allow fully coherent quantum operations over a moderate number of qubits.

References

[1] L. M. K. Vandersypen, M. Steffen, G. Breyta *et al.* Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance. *Nature (London)* **414**(2001), 883.

[2] T. Yamamoto, Y. A. Pashkin, O. Astafiev *et al.* Demonstration of conditional gate operation using superconducting charge qubits. *Nature (London)* **425**(2003), 941.

[3] D. Kielpinski, C. Monroe & D. J. Wineland. Architecture for a large-scale ion-trap quantum computer. *Nature (London)* **417**(2002), 709.

[4] O. Mandel, M. Greiner, A. Widera *et al.* Controlled collisions for multi-particle entanglement of optically trapped atoms. *Nature (London)* **425**(2003), 937.

[5] J.-M. Raimond, M. Brune & S. Haroche. Manipulating quantum entanglement with atoms and photons in a cavity. *Rev. Mod. Phys.* **73**(2001), 565.

[6] H. Mabuchi & A. C. Doherty. Cavity quantum electrodynamics: Coherence in context. *Science* **298**(2002), 1372.

[7] J. McKeever, A. Boca, A. D. Boozer *et al.* Experimental realization of a one-atom laser in the regime of strong coupling. *Nature (London)* **425**(2003), 268.

[8] A. Kuhn, M. Hennrich & G. Rempe. Deterministic single-photon source for distributed quantum networking. *Phys. Rev. Lett.* **89**(2002), 067901.
[9] S. Brattke, B. T. H. Varcoe & H. Walther. Generation of photon number states on demand via cavity quantum electrodynamics. *Phys. Rev. Lett.* **86**(2001), 3534.
[10] A. Rauschenbeutel, G. Nogues, S. Osnaghi *et al.* Step-by-step engineered multiparticle entanglement. *Science* **288**(2000), 2024.
[11] S. Haroche. Cavity quantum electrodynamics. In *Fundamental Systems in Quantum Optics, Les Houches Summer School, Session LIII*, J. Dalibard, J.-M. Raimond & J. Zinn-Justin, eds., North Holland, Amsterdam (1992). p. 767.
[12] D. Leibfried, R. Blatt, C. Monroe *et al.* Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75**(2003), 281.
[13] A. Einstein, B. Podolsky & N. Rosen. Can quantum mechanical description of physical reality be considered complete? *Phys. Rev.* **47**(1935), 777.
[14] X. Maître, E. Hagley, G. Nogues *et al.* Quantum memory with a single photon in a cavity. *Phys. Rev. Lett.* **79**(1997), 769.
[15] A. Rauschenbeutel, G. Nogues, S. Osnaghi *et al.* Coherent operation of a tunable quantum phase gate in cavity QED. *Phys. Rev. Lett.* **83**(1999), 5166.
[16] A. Rauschenbeutel, P. Bertet, S. Osnaghi *et al.* Controlled entanglement of two field modes in a cavity quantum electrodynamics experiment. *Phys. Rev. A* **64**(2001), 050301.
[17] E. Hagley, X. Maître, G. Nogues *et al.* Generation of Einstein-Podolsky-Rosen pairs of atoms. *Phys. Rev. Lett.* **79**(1997), 1.
[18] G. Nogues, A. Rauschenbeutel, S. Osnaghi *et al.* Seeing a single photon without destroying it. *Nature (London)* **400**(1999), 239.
[19] G. Nogues, A. Rauschenbeutel, S. Osnaghi *et al.* Measurement of a negative value for the Wigner function of radiation. *Phys. Rev. A* **62**(2000), 054101.
[20] P. Bertet, A. Auffeves, P. Maioli *et al.* Direct measurement of the Wigner function of a one-photon Fock state in a cavity. *Phys. Rev. Lett.* **89**(2002), 200402.
[21] M. Brune, S. Haroche, V. Lefèvre *et al.* Quantum non-demolition measurements of small photon numbers by Rydberg atoms phase sensitive detection. *Phys. Rev. Lett.* **65**(1990), 976.
[22] D. M. Greenberger, M. Horne & A. Zeilinger. Bell's theorem without inequalities. *Am. J. Phys.* **58**(1990), 1131.
[23] S. B. Zheng & G. C. Guo. Efficient scheme for two-atom entanglement and quantum information processing in cavity QED. *Phys. Rev. Lett.* **85**(2000), 2392.
[24] S. Osnaghi, P. Bertet, A. Auffeves *et al.* Coherent control of an atomic collision in a cavity. *Phys. Rev. Lett.* **87**(2001), 037902.
[25] D. Jacksh, J. I. Cirac, P. Zoller *et al.* Fast quantum gates for neutral atoms. *Phys. Rev. Lett.* **85**(2000), 2208.
[26] L. K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.* **79**(1997), 325.
[27] F. Yamaguchi, P. Milman, M. Brune *et al.* Quantum search with two atom collisions in cavity QED. *Phys. Rev. A* **66**(2002), 010302.
[28] P. Hyafil, J. Mozley, A. Perrin *et al.* Coherence-preserving trap architecture for long-term control of giant Rydberg atoms. *Phys. Rev. Lett.* **93**(2004), 103001.
[29] R. G. Hulet, E. S. Hilfer & D. Kleppner. Inhibited spontaneous emission by a Rydberg atom. *Phys. Rev. Lett.* **55**(1985), 2137.
[30] G. Gabrielse & H. Dehmelt. Observation of inhibited spontaneous emission. *Phys. Rev. Lett.* **55**(1985), 67.
[31] E. Peik. Electrodynamic trap for neutral atoms. *Eur. Phys. J. D* **6**(1999), 179.
[32] M. D. Lukin, M. Fleischhauer, R. Côté *et al.* Dipole blockade and quantum information processing in mesoscopic atomic ensembles. *Phys. Rev. Lett.* **87**(2001),

037901.
[33] R. Folman, P. Krüger, J. Schmiedmayer *et al.* Microscopic atom optics: from wires to atom chips. *Advances in Atomic, Molecular and Optical Physics* **48**(2002), 263.
[34] P. K. Day, H. G. LeDuc, B. A. Mazin *et al.* A broadband superconducting detector suitable for use in large arrays. *Nature (London)* **425**(2003), 817.
[35] A. Wallraff, D. Schuster, A. Blais *et al.* Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature (London)* **431**(2004), 162.
[36] L. Davidovich, N. Zagury, M. Brune *et al.* Teleportation of an unknown atomic state between two cavities using nonlocal microwave fields. *Phys. Rev. A* **50**(1994), R895.
[37] M. Brune, E. Hagley, J. Dreyer *et al.* Observing the progressive decoherence of the meter in a quantum measurement. *Phys. Rev. Lett.* **77**(1996), 4887.
[38] A. Auffeves, P. Maioli, T. Meunier *et al.* Entanglement of a mesoscopic field with an atom induced by photon graininess in a cavity. *Phys. Rev. Lett.* **91**(2003), 230405.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Possibility of Quantum Computation by Utilizing Carbon Nanotubes - Cooper Pair Splitting by Tomonaga-Luttinger Liquid -

J. Haruyama [a,b], K. Murakami [a], J. Mizubayashi [a], N. Kobayashi [a]
[a] *Aoyama Gakuin University, 5-10-1 Fuchinobe, Sagamihara, Kanagawa 229-8558 Japan*
[b] *JST-CREST, 4-1-8 Hon-machi, Kawaguchi, Saitama 332-0012 Japan*

Abstract. It is well known that carbon nanotubes (CNs) have strong spin coherence due to the unique electronic states. In addition, possibilities of the presence of strong spin entanglement in Cooper pairs and separation of them into individual spins by Tomonaga-Luttinger liquid states (TLLSs) in CNs have recently been reported. In the present study, we report 1. Evidence of spin splitting in Cooper pairs by the TLLSs in CNs and injection of individual spins into different CNs, 2. Finding of high-T_c intrinsic superconductivity competing to the TLLSs, and 3. The possibility of a molecular quantum spin entangler by utilizing them.

1. Introduction

The challenges of quantum computation have been investigated in many material systems, such as metal-superconductor and semiconductor quantum dots with single charging effect, flux-controlled superconductor rings, Bose-Einstein condensed atoms and molecules, and NMR. No studies of quantum computation, however, have been reported in carbon nanotubes (CNs), which are ideal one-dimensional (1D) molecular conductors and have exhibited a variety of exotic quantum phenomena. It is well known that CNs have strong electron spin coherence due to their unique electronic states and, hence, electron spins can be effectively transported via CNs without losing their phases over long distances and for long times [1]. This characteristic introduces unique quantum phenomena to CNs and provides the possibility of realizing quantum computation.

One realistic protocol in quantum computation, using CNs, is the creation of a quantum spin entangler (QSE) utilizing Cooper pairs injected from an s-wave superconductor. Indeed, two theoretical studies have predicted the possibility of QSE by proposing a hybrid system of a superconductor and two CNs within Tomonaga-Luttinger liquid states (TLLSs) and by utilizing separation of Cooper pairs by TLLSs [2,3], which are non-Fermi liquid states consisting of a repulsive

Coulomb interaction between electrons confined in 1D conductors, and have been frequently reported in CNs [4,5]. Strong repulsive electron-electron interaction in TLLS of CNs can efficiently separate spin singlets in Cooper pairs, which are injected from s-wave superconductors, into individual spins and inject them into different CNs, depending on tube spacing, dimension of superconductor, strength of TLLSs, etc.. Moreover, one experimental study reported the possible presence of strong spin entanglement of Cooper pairs in CNs by observing the co-existence of a Kondo singlet for the Kondo effect and a spin singlet in Cooper pairs in CN quantum dots [6]. However, there have been no other experimental studies to reveal the behaviors of coherence and entanglement in Cooper pairs of CNs, because it was difficult to efficiently inject Cooper pairs into CNs owing to poor interfaces of CNs/metal electrode junctions. In contrast, recently, we have successfully realized this by utilizing end-bonded CNs synthesized in nanopores of alumina templates and reported proximity-induced superconductivity in CNs [7]. In addition, we reported the possibility of Cooper pair separation into individual spins by TLLSs [14]. Here, in this paper, we report 1. Evidence of spin separation in Cooper pairs by the TLLSs and injection of the separated spins into different CNs, 2. Finding of high-T_c intrinsic superconductivity competing with the TLLSs, and 3. The possibility of a molecular spin entangler by utilizing them.

In Sec. 2, we will report that injected Cooper pairs are separated by the TLLSs following theories and individual spins exist in different CNs. Magnetic-field dependence in these spins is for the first time investigated. In Sec. 3, we report finding superconductivity with T_c as high as $12K$ realized by entirely end-bonding CNs. Interestingly, this allows observation of competition between coherence of emerged Cooper-pairs and TLLSs. In Sec. 4, we will propose a hybrid system of one superconductive CN and two Aharonov-Bohm ring CNs within TLLSs based on Secs. 3 and 4 and explain that this system can be a good candidate for a molecular spin qubit.

2. Splitting of Cooper Pair to Individual Spins by Tomonaga-Luttinger Liquids

Figure 1(a) shows the schematic cross-sectional view of our sample structures, i.e., an array of Au/Nb/multi-walled CNs (MWNTs)/Al prepared in an array of nanopores of alumina templates. We have reported a variety of quantum phenomena in CNs and metal nano-wires in this system [8]. This structure allows us to end-bond the MWNTs by evaporating a metal electrode on the top ends of the MWNTs (see (b) and (c)). This end-bonding has significantly enhanced the interface transparency of the metal electrode/CNs junctions. In Fig. 1, a Nb electrode has been used for this end-bonding with high temperature annealing because NbC is generated at the Nb/MWNTs interface, the transparency becomes higher.

The main panel of Fig. 2 shows the zero-bias resistance versus temperature relationship for different magnetic fields (H) in the sample with tube length ($L_{\rm tube}$) of $\sim 800nm$. The resistance starts to drop at $T = 3.8K$ (i.e., defined as $T_{\rm px}$) at $H = 0T$, due to the proximity effect of the Nb electrode. As the temperature decreases below $T = 1.2K$, which is the T_c of the Al substrate,

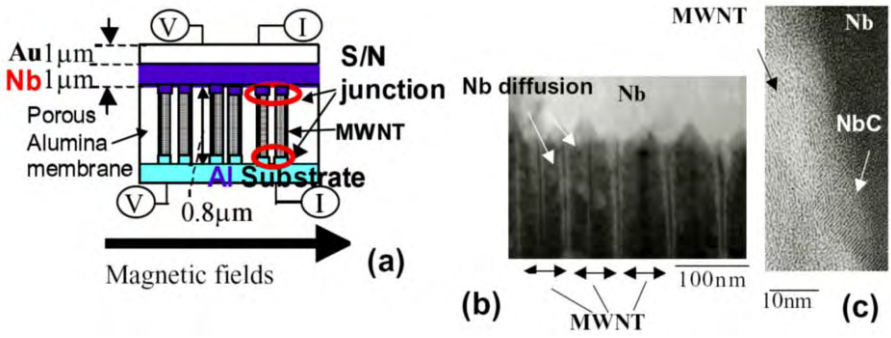


Figure 1. (a) Schematic cross-sections of an array of end-bonded Au/Nb/MWNTs/Al junctions, synthesized into the nanopores of alumina templates. T_c of our Nb was about $8.1K \sim 9K$ and H_c was about 1500 Gauss. The mean outer diameter of the MWNTs is about $80nm$, and the shell thickness was as large as $20nm$. (b) High-angle annular dark-field (HAADF) image of a cross-sectional (CS) TEM around the Nb/MWNT interface array. (c): High-resolution CSTEM images around the Nb/MWNT interface in (b)

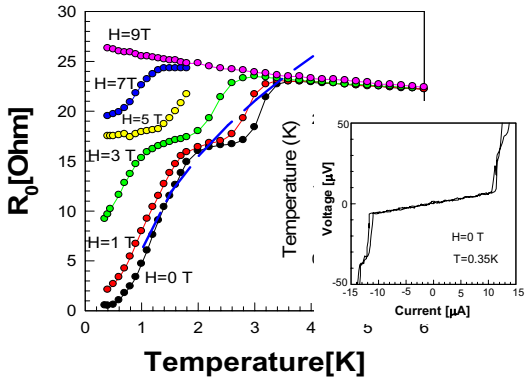


Figure 2. Zero-bias resistance (R_0) versus temperature relationship for different values of magnetic field perpendicular to the MWNT axis in the sample with $L_{\rm tube} =\sim 800nm$. The labels on the curves correspond to Tesla (T). The dotted line for the curve at H=0T is the result of the calculation obtained from the $R_0 \propto 14.3 \times T^{1/2}$ relationship Inset: Proximity-induced Josephson supercurrent.

the Cooper pair wave function (CPWF) starts to diffuse due to the proximity effect of Al. Consequently, one could successfully find a proximity induced superconductivity (PIS) at $T_c = 0.6K$ at $H = 0T$ and also a distinct proximity-induced supercurrent (inset). This presents strong evidence for the presence of highly-transparent Nb/MWNTs interfaces without any damage, and also highly-transparent MWNTs/Al interfaces.

The following behaviors are similar to those in single-walled CN (SWNT) junctions within a ballistic charge transport regime [9]. 1; The resistance gradually drops as the temperature decreases, showing two steps between the onset

temperature and T_c at $H = 0T$ (Fig. 2). 2; The transition between the superconducting state and the dissipative state is very abrupt, showing hysteresis loops and staircases at currents above I_c (inset).

The former can be understood by the pinning of the CPWFs, which rapidly diffuse into the MWNTs as temperature decreases. The latter, which is related to the inset, might be interpreted as a result of a (quantum) phase slip center, in which the CPWF is pinned by a normal conducting state formed in some parts of the superconductor wire and the phase slips throughout these parts. In contrast, there are some differences from the case of SWNTs. One of these differences is that the R_0 vs temperature relationship in Fig. 2 remains relatively unchanged, even when a field of $H = 1T$ was applied. Because the critical magnetic field (H_c) of Nb is about $0.15T$, the PIS should disappear under $H = 1T$. This anomalous behavior was discussed in ref. [7].

Figure 3 shows the doubly logarithmic scales of zero-bias conductance versus temperature relationship in the sample with $L_{\rm tube} = 600nm$ and $20nm$ tube spacing in a MWNT array. When the length of MWNTs for Fig. 2 was reduced below the mean free path ($\sim 700nm$) like this case, the MWNTs made a transition to a ballistic charge transport regime and exhibited a power law behavior at temperatures $> T_c \sim 10K$ (Fig. 3). Such power laws in the relationships of conductance vs. energies provide evidence for TLLSs in both multi-walled carbon nanotubes (MWNTs) [4] and SWNTs [5], with various powers α. The observed correlation exponent g, which denotes the strength of an electron-electron interaction (e.g., $g < 1$ and $g > 1$ signify the repulsive and attractive Coulomb interactions, respectively) and determines α, as low as ~ 0.2 in the previous studies stressed the presence of a strong repulsive Coulomb interaction in CNs.

Because the observed power exponent $\alpha =\sim 0.8$ at $\sim 10K < T <\sim 40K$ in Fig. 3(a) is in good agreement with α reported for Au-end junction [1], which just corresponds to our junction structure, this implies the presence of a TLLS. The Correlation exponent g can be estimated to be $g =\sim 0.25$ from the formula $\alpha = [1/g - 1]/4$ [4]. In contrast, this behavior disappears and conductance suddenly increases around T_c and, then, it starts to decrease again at $T < 7K$, following power law with $\alpha =\sim 1.28$ larger than the $\alpha =\sim 0.8$ mentioned above (Fig. 3(a)). This result stresses separation of Cooper pairs, which were injected from Nb, into individual spins by TLLSs of the MWNTs in accordance with the theory in [3].

Ref. [3] predicted that a tunneling probability Γ_{AB} of Cooper pairs from one superconductor to ends of two CNs within TLLSs is given by $\Gamma_{AB} =\sim (eV/h)(kT/\epsilon)[1/g - 1]/2$, when Cooper pairs are separated to two spins, which tunnel into two different CNs. We can obtain $\alpha =\sim 1.5$ by inserting $g =\sim 0.25$ into this formula. This $\alpha =\sim 1.5$ is approximately in agreement with the observed $\alpha =\sim 1.28$ in Fig. 3 and, hence, suggests a possible splitting of Cooper pairs by the TLLS. In order to reconfirm this argument, the results, which were measured in the samples with tube spacing of $30nm$ and $40nm$, are shown in Fig. 3 (b) and (c), respectively. In Fig. (c), the decrease in conductance, which appeared at $T < 7K$ in Fig. 3(a), entirely disappeared and only an increase in conductance (i.e., proximity-induced superconductivity) is observable down to $T = 1.5K$. In contrast, a decrease in conductance following power laws appears at $T < 3K$,

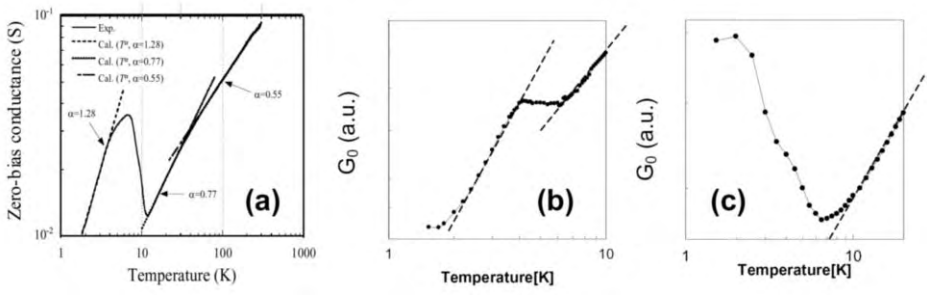


Figure 3. Temperature dependence of the zero-bias conductance G_0 on a doubly logarithmic scale in NbN/MWNTs/Al junctions for different tube spacing with the $600nm$-length MWNTs within a ballistic charge transport regime and TLL state. Tube spacings in arrays are (a) $20nm$, (b) $30nm$, and (c) $40nm$.

which is lower than $T = 7K$ in (a), in Fig. 3(b). These results are in qualitatively good agreement with the prediction by theory [2] as described below.

The theory [2] predicted the inequality $A(2\mu/\Delta)^{2\gamma} < (1/k_F\delta r)^{d-1}$ as the efficiency of Cooper pair splitting by TLLS in the hybrid system similar to ref. [3], where d and Δ are the dimension of the superconductor and superconducting gap, respectively. μ is the applied bias voltage between the superconductor and the leads. δr is the tube spacing, and γ is the exponent for tunneling into the bulk of a single TLLS, which is proportional to g. As this inequality is easily satisfied, the efficiency of Cooper pair splitting increases. This inequality implies that an increase in δr results in a lower probability in Cooper pair splitting. Because emergence of only PIS in Fig. 3(c) for the sample with the largest δr means no splitting of Cooper pairs, this agrees with the theory [2]. In addition, because the decrease in critical temperature for power law from $T = 7K$ in Fig. 3(a) to $T = 3K$ in Fig. 3(b) also implies reduction in probability of Cooper pair splitting, this qualitatively supports the theory [2].

Figure 4 shows the values of powers α in the low temperature region in Fig. 3(a) as a function of magnetic fields (H). The power laws still survive, even if $H = 7T$ is applied. The value of α is sensitive to H. We find that the value of α increases as H increases at $H <\sim 3T$ in any samples, while it monotonically decreases at $H >\sim 3T$, exhibiting small oscillatory behaviors except for sample 1.

Ref. [10] predicted that the value of α monotonically decreases as H increases in TLLS with a short range Coulomb interaction, because spins with opposite momenta localize at opposite ends of a CN due to the applied H and, hence, a repulsive Coulomb interaction between the electrons is reduced and the value of g increases, resulting in a decrease in α. This theory is in qualitatively good agreement with the decrease in α in the region of $H > 3T$ in Fig. 4. Hence, this can be evidence for the presence of TLLS. Because the tube length of about $600nm$ is large enough for a short-range interaction and Coulomb interaction is easily reduced by the edge localization, this agreement is appropriate.

In contrast, the increase in α at $H < 3T$ cannot be explained by this theory. The difference between the conventional TLLSs described by theory [10] and

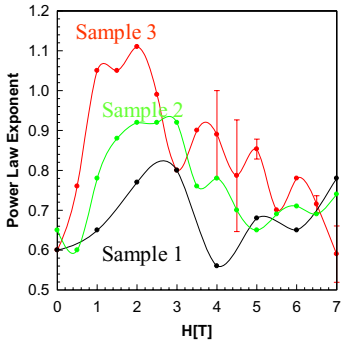


Figure 4. The value of power α in power laws at low temperatures in Fig. 3(a) as a function of magnetic field for different three samples.

our system is at least that separated Cooper pairs are injected into the ends of MWNTs in our case. Hence, this increase in α at low fields can be attributed to this difference. We propose one interpretation for this result as follows. When separated Cooper pairs are injected as individual spins into different MWNTs at $H = 0T$, the momentum of the spins are highly aligned in comparison to those in single spins injected from normal conductor electrodes. When two spins separated from Cooper pair are still in an entangled state, this alignment is much stronger. Moreover, MWNTs have strong spin coherence. Hence, when low fields are applied to such highly aligned spins, electron-electron interaction may be enhanced, resulting in a decrease in g and, hence, the increase in α as shown in Fig. 4. As the fields increase further and electron start to localize at both ends of the MWNT, this situation is destroyed and the value of α decreases. Therefore, this can be evidence for entangled spins existing in different MWNTs.

3. High-T_c Intrinsic Superconductivity in Entirely End-Bonded Nanotubes

The inequality $A(2\mu/\Delta)^{2\gamma} < (1/k_F\delta r)^{d-1}$ in ref. [2] also predicts that the efficiency of Cooper pair splitting can be enhanced by a reduction in the dimension of the superconductor (d) connected to the two different CNs. In this viewpoint, superconductive CNs gives the highest efficiency because they are within the $d = 1$ regime. Here, we report that the superconducting transition in MWNTs is realized by entirely end-bonding those top ends. Instead of the Nb electrode, we used just a normal conductor electrode (i.e., Au) for this investigation and entirely end-bonded the top ends of MWNTs by the AU electrode.

Figure 5(a) shows the result of zero-bias resistance (R_0) as a function of temperature. One can confirm an apparent superconducting transition with the onset T_c as high as $12K$ and the temperature ($T_c(R = 0)$), at where the R_0 drops to 0 Ω, as high as $T = 7.8K$. These values for onset T_c and $T_c(R = 0)$ are at least approximately 30- and 40-times, respectively, higher than those reported in SWNT ropes [11]. Figure 5(b) shows the differential resistance as a function of the current for different temperatures. A low and broad resistance peak due to

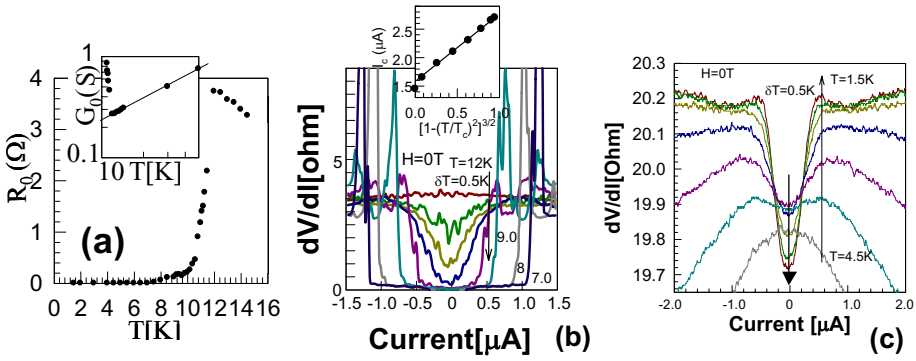


Figure 5. (a) Zero-bias resistance as a function of the temperatures in the entirely end-bonded MWNTs with N = 9 (b) Differential resistance as a function of the current for the different temperatures in sample (a). Inset: Critical current vs normalized temperatures for GL theory. (c) Differential resistance as a function of the current for different temperatures in partially end-bonded MWNTs with $1 < N < 9$.

the TLLS exists at $T = 12K$. This peak disappears suddenly and a resistance dip appears $T = 11.5K$. The depth and width of this resistance dip monotonically increase as the temperature decreases, corresponding to Fig. 5(a). The value of superconducting gap $\Delta \approx 1.15meV$ observed at $T = 8K$ in this sample is in excellent agreement with the Bardeen-Cooper-Schrieffer (BCS) relation $\Delta = 1.76kT_c$, when $T_c(R = 0) = 7.8K$ is used. Moreover, the behaviors of critical current (I_c) in normalized temperatures as shown in the inset of (a) is also in excellent agreement with the Ginzburg-Landau (GL) critical current behavior for a homogeneous order parameter, $I_c \propto [1-(T/T_c)^2]^{3/2}$ [12]. These suggest that the observed superconductivity is strongly associated with the BCS type. One of the most interesting points in this superconductivity is the discovery of competition between the TLLS and superconductive states, which strongly depends on the number of electrically active shells (N) in MWNTs. For instance, superconducting transition cannot appear and only the TLLS (i.e., conductance decrease on temperature decrease following power laws and corresponding R_0 peak) is observable in the case of $N = 1$, which was formed by Au electrode bulk MWNT junction. In contrast, Fig. 5 corresponds to $N = 9$, because of the entire end-bonding. As shown in Fig. 5(b), a resistance peak, which is observable at $T = 12K$ and means presence of the TLLS, suddenly disappears and a resistance dip, which corresponds to superconducting gap, appears at $T = 11.5K$ Hence, the strength of the TLLS is weak and the superconducting transition can easily overcome the TLLS in this system.

Figure 5 (c) shows the differential resistance as a function of the current for different temperatures in the sample with $1 < N < 9$. This is very different from Fig. 5(b). A large and broad resistance peak due to the TLLS is observable at $T = 4.5K$. It broadens as the temperature decreases and mostly disappears at $T = 2.5K$. In contrast, a resistance dip with a narrow width due to superconductivity begins to appear at the center of the peak at $T = 4K$ and the depth monotonically deepens as the temperature decreases. A corresponding gradual dip in

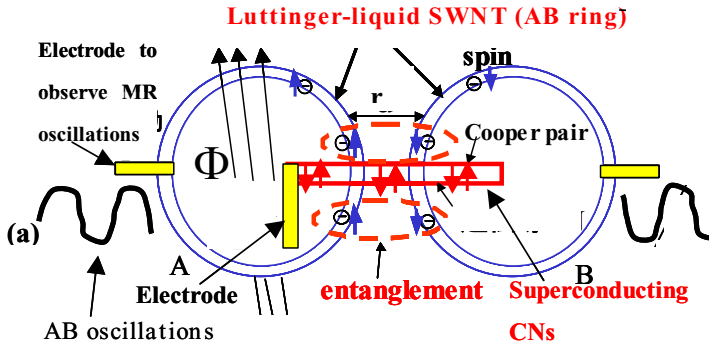


Figure 6. An example of molecular quantum spin entangler. A hybrid system of one superconductive CNs and two Aharonov-Bohm (AB) ring CNs within TLLSs.

the total R_0 can be observed only below $T = 3.5K$, reflecting the superposition of the resistance peak and drop. Importantly, the presence of this resistance peak prevents both the emergence of superconducting gap at $T \geq 4.5K$ and the R_0 drop at $T \geq 3.5K$. These results obviously indicate a competition between the TLLS and superconductivity. Such a relationship of the TLLS and a short range attractive Coulomb interaction via phonons depending on N has been predicted in a rope of SWNTs by Gonzalez [13]. One can interpret our results by replacing the number of SWNTs included in a rope with N in a MWNT.

This competition between the TLLS and superconductivity is very different from the case of PIS (i.e., those in Fig. 3) and, hence, interesting. Anyway, one can enhance the efficiency of Cooper pair splitting by utilizing this superconductive MWNT.

4. Proposal of a Molecular Quantum Spin Entangler

Here, Fig. 6 shows one of examples of a molecular QSE, based on a hybrid system of one superconductive CN and two normal conductive CNs within TLLSs discussed above. Normal conductive CNs form Aharonov-Bohm (AB) rings in this system. Cooper pairs are efficiently separated to individual spins by TLLS in the AB-ring CNs, because the superconductor CN is within $d = 1$ regime, and the separated spins are injected to different AB rings with opposite spin momenta. As shown in Fig. 4, the electron-electron interaction is enhanced at low fields. Hence, when a low magnetic field is applied so as to penetrate only into one AB ring, AB oscillation can be very sensitive to the applied fields. If two spins existing in different AB rings are still in an entangled state, the spin in the other AB ring should exhibit opposite-phase AB oscillation even without applying magnetic fields. Therefore, one can investigate spin entanglement by simultaneously observing these AB oscillations. If Cooper pairs in CNs actually have a strong entanglement [6,14], separated spins retain a strongly entangled state crossing the different CNs. This system results in realization of good molecular QSE.

5. Conclusion

We have reported 1. Evidence of efficient spin splitting in Cooper pairs, which were injected from Nb electrode, by the TLLSs in the MWNTs and injection of the separated spins into different MWNTs and 2. Finding of $T_c = 12K$ intrinsic superconductivity in entirely end-bonded MWNTs, which overcomes the TLLSs, and proposed 3. The possibility of a molecular quantum spin entangler by utilizing these two terms. No study for realizing quantum computation by utilizing CNs has been reported to date. Only we are approaching to this realization based on the present results. CNs can be good candidates for realizing molecular quantum computation with strong spin coherence and entanglement.

Acknowledgements We are grateful to P.Recher, T.Akazaki, H.Takayanagi, H.Shinohara, S.Maruyama, R.Saito, S.Saito, S.Tarucha, Y.Iye, W.Tsukada, D.Loss, J.-P. Leburton, M.Dresselhaus, and H.Bouchiat for fruitful help, discussion, and encouragement.

References

[1] L.Langer, V.Bayot, J.-P.Issi, et al., Phys.Rev.Lett **76**, 479 (1996)
[2] P.Recher and D. Loss, Phys. Rev. B **65**, 165327 (2002)
[3] C.Bena et al., Phys.Rev.Lett. **89**, 037901 (2002)
[4] A.Bachtold, C.Shonenberger, et al., Phys.Rev.Lett. **87**, 166801 (2001); M.R.Buitelaar, et al., Phys.Rev. Lett. **88**, 156801 (2002); A.Bachtold, et al., Nature **397**, 673 (1999)
[5] M.Bockrath, et al., Nature **397**, 598 (1999); H.Ishii, H.Kataura, et al., Nature **426**, 540 (2003)
[6] M.R.Buitelaar, T.Nussbaurner, C.Shonenberger, Phys.Rev. Lett. **89**, 256801 (2002)
[7] J.Haruyama, et al., Phys.Rev.B 68, 165420 (2003); Appl.Phys.Lett. 84, 4714 (2004); Microelectronics Journal 34, 537 (2003); Physica C 408, 85 (2004)
[8] J. Haruyama, et al., Phys. Rev. B **65**, 33402 (2002); Phys.Rev.B, 073406 (2001); Appl.Phys.Lett. **81**, 3031 (2002); Appl.Phys.Lett. **77**, 2891 (2000)
[9] A. Yu. Kasumov, H. Bouchiat, et al., Science **284**, 1508 (1999); Phys.Rev.B **68**, 214521 (2003)
[10] S.Bellucci and P.Onorato, Cond-mat/0504170
[11] M. Kociak, H. Bouchiat, et al., Phys. Rev. Lett. **86**, 2416 (2001)
[12] M.Tinkam, Introduction to Superconductivity (McGraw-Hill, New York 1996)
[13] J.Gonzalez, Phys.Rev.Lett. **88**, 076403 (2002); A.Sedeki, et al., Phys.Rev.B **65**, 140515(R) (2002); J.Gonzalez, Phys.Rev.Lett. **87**, 136401 (2001)
[14] J.Haruyama, et al., Phys.Rev.Lett. In submission; Physica Stat. Sol. (b) 242(2), 265 (2005)

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Implementing Quantum Processors in the Solid State

Crispin H. W. Barnes

Abstract. This paper illustrates the themes in my three NATO school talks by discussing two simple modifications to the electron-spin-based surface-acoustic-wave quantum-state processor: the use of variable g-factor materials is introduced as a means to have electrical control over single-qubit rotation; and a form of optical input and readout that relies on the coherent creation and recombination of entangled excitons is introduced as a means to allow this processor to be used as a component in a quantum communications network.

Keywords. Solid State, Quantum Processor,

1. Introduction

An N-qubit quantum processor is capable of storing and manipulating vectors with 2^N elements. This vast capacity for computation and manipulation has captured the attention of academics, industry and government. These communities are now actively engaged in an attempt to identify N-particle systems that not only satisfy the DiVincenzo criteria but which will also minimize the time and cost of production of even a single quantum processor. This paper suggests two refinements to the surface acoustic wave (SAW) quantum computer ('Barnes *et al.* 2000' [1]) that make it faster, easier to control, easier to program and easier to read-out the final state. In addition, these modifications make it more attractive for use as a component in a quantum communication network because both the input and the output can be in the form of entangled photons.

The original design for SAW quantum processor consists of a set of N depleted, parallel channels. The magnetic field along each of the channels is modulated by a pattern of single-domain magnetic surface gates of different orientations, lengths and field strengths. Electrons can tunnel between channels via electrostatically controlled breaks in the barriers between channels. N electrons are captured in each cycle of the SAW so that each SAW minimum has one electron trapped in each of the parallel channels. Each set of N electrons in each SAW minimum is a single quantum processor. Each of these quantum processors is put into the same initial state - through the physics of the SAW electron capture process [1],[2]. The state of each processor is manipulated in exactly the same way as it passes through single-qubit gates, produced by the magnetic gates, and the two-qubit gates, defined by the electrostatic tunnel barriers. Measurement of each qubit occurs at the end of each channel and the average over a large number of measurements of the same qubit in successive processors is the output of the processor device. This measurement is the same as the ensemble measurement made in NMR quantum computation where a large number of identical systems are prepared, manipulated and then measured. The most important difference between a SAW quantum processor and

an NMR quantum processor is that owing to its low temperature operation the initial states are likely to be nearly identical and highly pure.

The modifications that are proposed and discussed in this paper are: the use of variable g-factor materials for single qubit manipulation; and the use optical input and readout.

2. Variable g-factor Materials

Many solid-state approaches to producing a quantum processor use electron spin as the qubit. In particular, this is the case for the original SAW quantum processor [1]. Within its architecture, single-qubit operations are performed by placing split, single-domain magnetic gates across a channel, or perpendicular to it, at the location where the rotation is required. Since only rotation about two perpendicular axes is required to move a spin to an arbitrary point on its Bloch sphere, this combination is sufficient to produce universal single-qubit gates. The main problems with such gates are that: fairly large magnetic fields are required if the length of each magnetic gate is not to be longer than the spin dephasing length (typical gate lengths in GaAs, for 1T surface fields, exceed 1μm); once the magnetic gates have been patterned, there is no easy way to change their magnetic field strength, especially if there is a distributed pattern of similar magnetic gates nearby; each single-domain magnet would produce unwanted stray fields affecting other channels and the orientation of domains in other magnets; and owing to the nature of their fabrication, evaporation through a mask, it is very difficult to determine what field strength will be achieved for a magnetic gate produced using a given mask. A degree of control over single qubit rotation can be achieved if electromagnets are used in place of permanent magnets but in order to achieve sufficient field strengths very high currents (1A) in very short pulses (1ms) are required.

The use of variable g-factor materials solves all of these problems. The original proposal suggesting the use of such materials by Vrijen *et al* [3] consisted of a SiGe heterostructure containing a single quantum well that was split into two parts with different g-factors. This is possible because there are two types of SiGe with different Ge content that have the same band gap, band offset and are lattice matched but have different g-factors. The doping in their heterostructure is arranged so that the two-dimensional subband wavefunction sits entirely in one half of the quantum well. A continuous-wave off-resonance global r.f signal and a constant external magnetic field are applied to the device. When a single-qubit operation is required a potential is applied to a surface gate above the qubit so that the electron is pulled to the other side of the quantum well where it comes onto resonance with the external r.f. field. σ_x and σ_y rotations are achieved by carefully timing the gate-potential pulses with the precession of the electron spin in the external field. A very similar scheme is possible in III-IV materials. There are two basic options: create a double quantum well structure in which the upper and lower wells have different Aluminium or Indium content; or create a single quantum well with a graded barrier of Aluminium. Aluminium content tends to reduce the g factor towards and through zero. Indium content tends to increase the g factor and can do so by roughly an order of magnitude.

In order to apply these variable g-factor ideas to the SAW quantum processor one must either use the SiGe heterostructures and use flip-chip technology to bind them to a

piezoelectric substrate such as quartz [4] or use one of the two III-IV options. σ_x and σ_y rotations are achieved by placing the surface gates distances apart that correlate with the precession of the electron spins, in a constant external magnetic field, with their position in the device as they are moved through by the SAW.

3. Two-Gubit Gate

The two-qubit operation for electron spins in a SAW quantum computer is based on the exchange interaction. It occurs when two electrons are brought into tunnel contact by reducing a barrier separating them. For a SAW quantum processor device the barrier is defined electrostatically by a short metallic Schottky gate placed over a gap in the barrier between channels. These gates are typically an order of magnitude faster than the single-qubit magnetic gates. The necessary gate length for root-of-swap operations can be tens of nanometers. The effect of each gate can be made very local since the tunnelling probability is exponentially dependent on surface gate voltage. This implies that their operation is unlikely to affect the operation of the single qubit gates. On the contrary though, the potential applied to a single quit gate could change an adjacent tunnel barrier. Therefore, in order to define a set of single and two-qubit gates using variable g-factor materials a process of calibration will need to be carried out to define any quantum algorithm.

4. Optical Readout

From the GaAs optical selection rules a spin +1/2 electron will recombine with a spin +3/2 hole to produce a left-handed circularly polarized photon ($m_j = -1$) ('D'yakonov *et al.* 1971' [5],'Pierce *et al.* 1975' [6],'Vrijen *et al.* 2000' [4]). Similarly a spin -1/2 electron will recombine with a spin -3/2 hole to produce a right-handed circularly polarized photon. Optical readout of SAW electron-spin qubits can therefore be achieved by bringing the SAW electrons into contact with holes at the end of each channel and measuring the polarization of the output light. This can be done in two ways: placing a hole gas at the end of each channel ('Foden *et al.* 2000' [7],'North *et al.* 1993' [8], 'Vijendran *et al.* 1999' [9]) or, at the input stage, create excitons so that each electron carries a hole along with it('Sogawa *et al.* 2001' [10]).

For the hole gas case, as each electron enters the recombination region it will be decohered by elastic collisions with holes so that its state will be spin up with probability $|\alpha|^2$ and spin down state with probability $|\beta|^2$. The two different hole species occur with equal probability and therefore the numbers of left and right circularly polarized photons emitted will be in direct relation to these probabilities. The output light from each channel is passed through a beam splitter and perpendicular polarizing elements to measure the ratio $|\alpha|^2 : |\beta|^2$. This optical readout method does not transfer the quantum state of the entangled electrons to the emitted photon field. For quantum communication though, accurate quantum information transfer is clearly essential. Vrijen *et al.* (2000) [4] has suggested two methods to achieve coherent quantum information transfer. They both require placing each hole gas in a high finesse cavity and straining the GaAs or applying an external magnetic field to remove the valance-band degeneracy. A suitable

high-finesse cavity can be produced during molecular beam epitaxy growth. Alternating layers of material with different dielectric constants are grown in bands above and below the hole gas. These layers then act as Bragg reflectors. The introduction of strain is possible through introducing layers of material with different lattice constants. Imamoglu *et al.*(1999)[11] has suggested and alternate method for preserving entanglement that involves the use of conduction-band-hole Raman transitions induced by classical laser fields but I will not discuss that here.

5. Optical Input

When a single photon hits GaAs-AlGaAs heterostructure it is likely that it will create an exciton. If the photon is of the appropriate polarization and the heterostructure has the degeneracy of it hole bands lifted, the quantum state of the incident photon will be transferred to the electron part of the exciton wavefunction and the electron and hole parts of the exciton wavefunction will be separable [4]. If there is a SAW present then both the electron and the hole will be swept along but kept physically separate. The electron will move to a SAW minimum and the hole to a maximum. The electron can then be guided electrostatically into a narrow channel where its spin can be manipulated and coherently re-emitted by the methods described above. With this method of coherent optical input, the recombination can be with the original hole created with the electron or another, since by design, the electron and hole quantum information are separable.

The extension of this idea to N-photons is straight-forward but does not scale especially well. It would be necessary to arrange for each photon to arrive at the device in a different optical fibre so that the electron from each exciton created can be forced into a different channel. If delays occur, it is possible that different electrons may be captured in different SAW minima. This could be compensated for by using quantum dots to capture and delay electrons until all N have been captured or an error registered.

6. The Optical Quantum-State Processor

A SAW-based N-qubit optical quantum-state processor would require N input optical fibres connected to GaAs regions where, by using a suitable cavity, the probability for the creation of a single exciton for each incident photon would be above some predefined error tolerance threshold. These collection regions would connect directly to N depleted parallel channels. A regular pattern of surface gates defining both single and two qubit operations would have potentials applied to them to carry out a specific quantum algorithm on the electron part of the N captured excitons. At the end of the device suitable screening gates, cavities and optical fibres would be attached so that recombination would create an output set of entangled photons.

A single computation would be performed by sending a set of N entangled photons along the attached optical fibres. They would create a set of N excitons, with an entangled electron part to the full many-particle wavefunction but with the electron and hole parts separable. These electrons would be swept through the N channels by a SAW wave. Delay methods using quantum dots would be performed if necessary. Single qubit operations would be performed by electron-spin-resonance in regions where gates were

placed in the center of channels. Two qubit operations would be performed where adjacent electrons were allowed into tunnel contact. At the end of the device recombination, would cause the N photons to be re-emitted in a different state of entanglement.

A simple example of the processor described could consist of two photons creating two excitons, the electron parts of which are captured in a single SAW minimum and swept in a single depleted channel. Over the length of the channel, the two electrons will interact strongly and become entangled. Decoherence would cause them to fall into a singlet state. At the end of the channel a Y-branch splitter is used to separate the two electrons. They are then passed into two recombination regions, where they create two entangled photons.

7. Summary

This paper has shown how concepts from the NATO school talk can be used to modify the SAW quantum computer proposed by Barnes *et al.* (2000) [1]. Single-qubit gates based on variable g factor materials and electron-spin-resonance techniques were introduced. They have the advantage over the originally proposed magnetic split gates that they can be tuned simply by changing the potential on a surface Schottky gate. Optical input and output were also described and it was shown how a SAW processor could be used as a non-linear element, modifying the state of entanglement of a set of N photons, in an optical network. The important issues of error production and decoherence for this type of quantum computer are discussed in Barnes *et al.* (2000) [1].

Acknowledgements: I acknowledge funding from the EPSRC and useful conversations with Chris Ford.

References

[1] Barnes, C. H. W., Shilton, J. M. , and Robinson, A. M. , Phys. Rev. B 62, 8410 (2000).
[2] Robinson, A. M., Barnes, C. H. W. Phys. Rev. B 63 165418 (2001).
[3] R. Vrijen R., Yablonovitch E.,Wang K. L. ,Jiang H. W. , Balandin A. ,Roychowdhury V., Phys. Rev. A 62, 12306 (2000).
[4] Vrijen R, Yablonovitch E, PHYSICA E 10 (4): 569-575 2001
[5] D'yakanov, M. I. and Perel, V. I., Soviet Phys. JETP 33 1053 (1970).
[6] Pierce, D. T. Meier, F. and Zurcher, P. Phys. Letts. A 51 465 (1975).
[7] Foden, C. L. ,Talyanskii, V. I. ,Milburn, G. J. ,Leadbeater, M. L. ,Pepper, M. Phys. Rev. A 62 011803 (2000).
[8] North, A. J., Burroughes, J. H., Burke, T., Shields, A. J., Norman, C. E., and Pepper, M., IEEE J. Quantum Electron. 35 352 (1999).
[9] Vijendran S, Sazio PJA, Beere HE, Jones GAC, Ritchie DA, Norman CE J. Vac. Sci B 17 3226 (1999).
[10] Sogawa, T. ,Santos, P.V. ,Zhang, S.K. ,Eshlaghi, S. ,Wieck, A. D. , and Ploog, K. H. 87 276601-1 (2001)
[11] Imamoglu, A. Awschalom, D. D., Burkard, G., DiVincenzo, D. P., Loss, D., Sherwin, M. and Small, A. Phys. Rev. Lett 83 4204-4207 (1999).

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Quantum Field Trajectories Under Photon Number QND Measurement

Alexander A. Bukach, Sergei Ya. Kilin
B. I. Stepanov Institute of Physics, NAS Belarus,
Nezavisimosti ave., 68, 220072 Minsk, Belarus

Abstract. We study quantum stochastic trajectories of a field under quantum non-demolition measurement of its photon number using a scheme based on the optical Kerr effect. The signal field is shown to evolve into a state consisting of two Fock states (Fock state qubit) entangled with a probe, with the lifetime of such a state being much longer than the mean time between counts on the detectors. However, this pairwise state is not stable, excepting some particular cases, and it tends asymptotically with time to a pure Fock state.

Keywords. Quantum non-demolition measurement, continuous measurement, Fock states qubit, Fock state.

1. Introduction

Quantum non-demolition (QND) measurements are very important for the investigation of optical fields [1]. The idea of the QND first appeared in the paper by Braginsky and Vorontsov, who proposed a way to perform the measurement without feedback on the measured observable [2] and called such a measurement a QND. Soon afterwards, the idea of QND measurements was applied to problems of quantum optics. A number of methods for QND measurements of a field's photon number and quadratures were suggested, one of which is considered here as the photon number QND measurement based on the optical Kerr effect [3].

Let us recall that the refractive index of a Kerr medium changes linearly with the intensity of the field passing though it. Therefore, if we aim a signal field a at such a medium, it changes its refractive index, and once we let a probe field c pass through the medium, it gets a phase shift depending on the intensity of the signal. It is this fact that is used in the QND measurement scheme suggested by Imoto *et al.* [3] (Fig. 1). The interaction of the fields in the Kerr medium is described by the Hamiltonian:

$$H_I = \hbar\kappa a^+ a c^+ c, \tag{1}$$

where κ characterizes the nonlinearity of the medium. Since the signal photon number operator $n_a = a^+ a$ commutes with this Hamiltonian, then n_a is an integral of motion during the interaction, which is a necessary condition for the measurement to be QND [1].

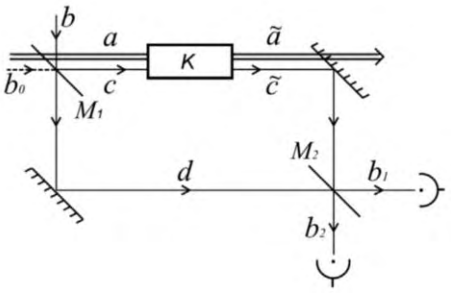


Figure 1. Photon number QND measurement scheme. The probe mode b mixing at the semi-transparent mirror M_1 with the vacuum mode b_0 is divided onto two beams: c and d. The former interacts with the signal field a in a nonlinear Kerr medium K. Field modes $\tilde{c}$ and d are aimed at the detectors of a homodyne scheme, which measures probe phase shift being proportionate to the signal photon number.

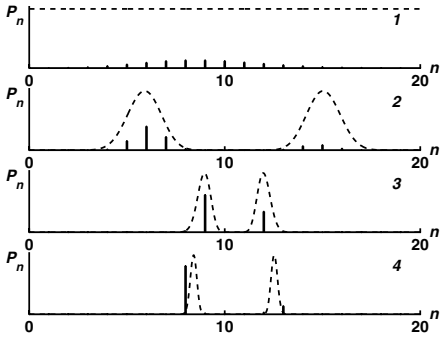


Figure 2. An example of the transformation of the signal photon number distribution during the measurement. The distribution is obtained by multiplying the initial distribution by the selection curve shown by the dashed line. *1)* 0 counts, *2)* 15 counts, *3)* 100 counts, *4)* 400 counts.

The phase shift providing the information on the signal photon number can be measured here by means of the homodyne detection of the probe. At that the measurement of the photocurrent in the detector is usually considered as a momentary act which is equivalent (to a detector quantum efficiency) to the measurement of the photon number in the corresponding probe mode [1], which in turn defines the phase shift and the signal photon number n_a. Therefore, the measurement of photocurrents difference seems to reduce the signal to the corresponding Fock state. However, in fact, the current is nothing but a sequence of photocounts, and "momentary" consideration is correct only at times much greater than mean time between counts $\bar{\tau}$. If one considers a time comparable to $\bar{\tau}$ then the "reduction" process appears to take some finite time, and there is evolution rather than reduction taking place during the measurement. It is this evolution that is the issue of investigation of the present work. Using the techniques of the theory of quantum continuous measurement [5], we investigate how the fields evolve with time from an initial to a final Fock state [6].

2. Quantum Field Trajectories

The evolution of the density matrix ρ of the fields interacting with the detectors is described by a Master equation that, in Lindblad form, can be written as [4]:

$$\dot{\rho} = -\frac{i}{\hbar}[H_f, \rho] + \frac{\gamma}{2}\sum_{l=1}^{2}([b_l\rho, b_l^+] + [b_l, \rho b_l^+]), \tag{2}$$

where γ is some constant that depends on the properties of the detectors and defines the efficiency of their interaction with the field. The fields Hamiltonian $H_f = \hbar\omega_a a^+ a + \hbar\omega_p \sum_{l=1}^{2} b_l^+ b_l + V(t)$ includes free Hamiltonians of both the

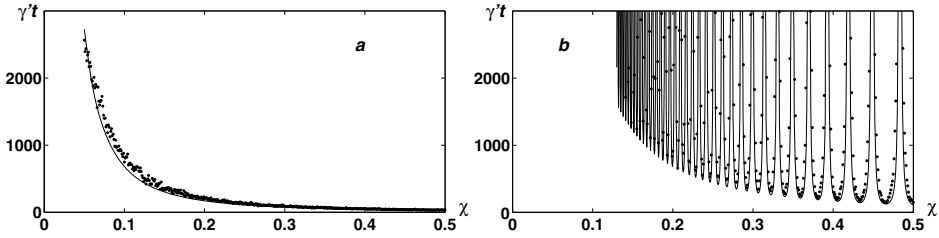


Figure 3. Mean lifetime of the pair of Fock states. Plots a and b correspond to the cases of different width of the initial distribution (one and two peaks of selection function over its width, respectively). The results of numerical simulation are represented by dots, time estimations are shown by solid lines.

signal and probe fields with the frequencies ω_a and ω_p respectively, as well as Hamiltonian of excitation of the probe field $V(t) = \hbar(v^* e^{i\omega_0 t} b + v e^{-i\omega_0 t} b^+)$ at the frequency of ω_0.

The solution of Eq. (2) does not provide perfect information on the field state since it is averaged over an ensemble of the measurements and thus does not take into account the information on registered photocounts. To describe the state of the fields in one particular run of the measurement, one can rewrite the averaged density matrix $\rho(t)$ as a linear combination of operators corresponding to different sequences of counts on the detectors (so-called conditional density matrices) [5].

Having calculated these conditional density matrices, we analyzed stochastic trajectories of the signal field, i. e. the evolution of the field under the registration of a random sequence of counts by the detectors, and also performed a numerical simulation of the trajectories (Fig. 4). The photon number distribution of the signal during the measurement is defined as a product of an initial distribution by some "selection" function, the plot of which is just the sequence of peaks with the width decreasing with time. Hence, during the registration of counts there appear regions in the distribution in which the probability of finding photons grows while outside it vanishes (Fig. 2). After some time, there are only a few nonzero probabilities, and finally only a single probability stays in the distribution, i. e. the signal field reduces to a Fock state (Fig. 4.8). At the same time, the phase distribution of the signal becomes wider, tending to a homogeneous one. The photon number in the final Fock state of the signal proves to be defined by the difference of mean photocurrents in the detectors.

A specific feature of the field evolution is the existence of an intermediate state having only two nonzero probabilities in its photon number distribution (Fig. 4.4 – 4.6). It thus consists of two Fock states of the signal field (so-called Fock state qubit) entangled with the probe. Such a state appears to exist for most of the time of a reduction of the initial state to a final Fock state. Its lifetime can be estimated analytically and proves to be much longer than the mean time between counts $\bar{\tau}$ (Fig. 3). This fact allows us to consider the scheme of QND measurement used here as a way of generating such pairwise states.

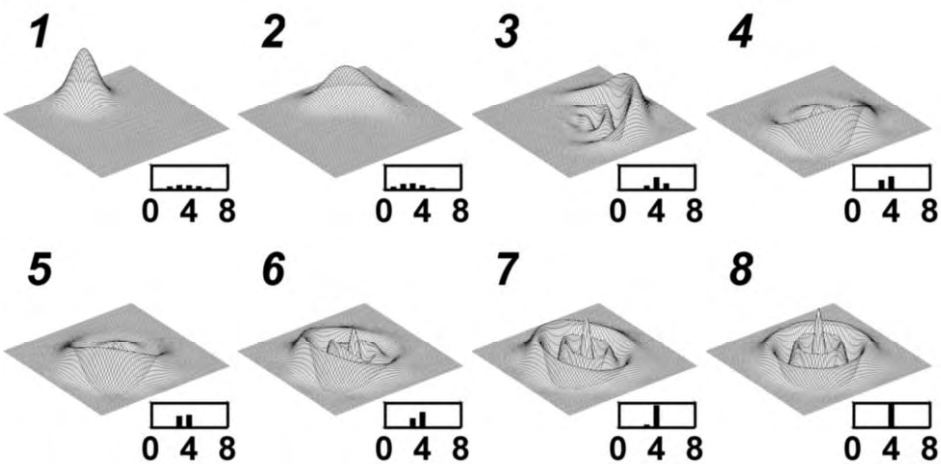


Figure 4. The Wigner function and the photon number distribution (insets) of the signal field at the subsequent points of time at one particular run of the measurement (numerical simulation).

3. Conclusions

In this paper, we have studied the evolution of the (signal) field undergoing a QND measurement based on its interaction with the probe field in a nonlinear Kerr medium, by applying the theory of continuous measurement. It was shown that in any particular run of measurement, the signal field reduces to the state being a superposition of two Fock states (Fock state qubit) entangled with the probe, with the lifetime of this pairwise state being much longer than the mean time between counts in the photo-detection scheme. However, these states are not stable except in some particular cases, and they tend asymptotically with time to a single Fock state.

Acknowledgements

The authors wish to thank the Basic Research Foundation of Belarus and INTAS fund for partial support of the present work. We are also indebted to Dr. Dmitri B. Horoshko for his useful discussions.

References

[1] M. O. Scully and M. S. Zubairy, *Quantum optics*, Cambridge Univ. Press (1999).
[2] V. B. Braginsky, Yu. I. Vorontsov, Uspekhi Fizicheskih Nauk **114**, 41 (1974).
[3] N. Imoto, H. A. Haus, Y. Yamamoto, Phys. Rev. A **32**, 2287 (1985).
[4] H. D. Carmichael, *An Open System Approach to Quantum Optics*, Springer-Verlag, Berlin (1993).
[5] S. Ya. Kilin, *Quantum optics: fields and their detection*, Nauka i tekhnika, Minsk (1990) (in Russian).
[6] A. A. Bukach, S. Ya. Kilin, submitted to JETP.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Quantum Computation Beyond the "Standard Circuit Model"

K. Ch. Chatzisavvas [a,1], C. Daskaloyannis [b] and C. P. Panos [a]
[a] *Department of Theoretical Physics, Aristotle University of Thessaloniki 54124 Thessaloniki, Greece*
[b] *Department of Mathematics, Aristotle University of Thessaloniki 54124 Thessaloniki, Greece*

Abstract. Construction of explicit quantum circuits follows the notion of the "standard circuit model" introduced in the solid and profound analysis of elementary gates providing quantum computation. Nevertheless the model is not always optimal (e.g. concerning the number of computational steps) and it neglects physical systems which cannot follow the "standard circuit model" analysis. We propose a computational scheme which overcomes the notion of the transposition from classical circuits providing a computation scheme with the least possible number of Hamiltonians in order to minimize the physical resources needed to perform quantum computation and to succeed a minimization of the computational procedure (minimizing the number of computational steps needed to perform an arbitrary unitary transformation). It is a general scheme of construction, independent of the specific system used for the implementation of the quantum computer. The open problem of controllability in Lie groups is directly related and rises to prominence in an effort to perform universal quantum computation.

Keywords. Quantum Gates, Quantum Computation, Quantum Control Theory

1. The "Standard Circuit Model"

The "standard circuit model" is an established proposal to implement quantum gates in quantum computation [1]. In this model essential is the notion of the universal gate [2]. Thus, any given quantum gate (any given unitary transformation of the quantum system that implements the quantum computer) can be analyzed using a set of basic gates, known as universal gates. The selection of the set of universal gates is not unique [3]. One-qubit gates can be analyzed using only Hadamard and phase gates. Two-qubit gates can be analyzed using Hadamard, phase and the CNOT gate and this is generalized in the case of N-qubit gates, while it was noted that in the general case an infinite number of steps are needed to perform a gate explicitly [4].

In the "standard circuit model", physical systems are neglected if they cannot copy easily the model (if someone cannot perform easily one of the selected universal gates). Also, neither the number of computational steps nor the total time to perform computation are optimal [5].

[1]Correspondence to: K. Ch. Chatzisavvas, Department of Theoretical Physics, Aristotle University of Thessaloniki, Greece. Tel.: +30 231 0998234; Fax: +30 231 0998128; E-mail: kchatz@auth.gr.

2. Quantum Control Theory

In quantum control theory, the generalization of the control theory in quantum systems, a system is said to be controllable if an arbitrary Lie group element $W \in SU(2^N)$ can be decomposed in finite time as

$$W = e^{-ia_n J^{(n)}} \ldots e^{-ia_2 J^{(2)}} e^{-ia_1 J^{(1)}} \tag{1}$$

where $J^{(k)} \in \{J_1, J_2, \ldots, J_m\}$ are generators of the corresponding $su(2^N)$ Lie algebra and $a_i \in \mathbb{R}$. In the case of quantum computation, W is equivalent with an arbitrary unitary transformation (up to a global phase) so it is equivalent with an arbitrary N-qubit gate. $J^{(k)}$ corresponds to the Hamiltonians describing the system under consideration while a_i are equivalent with time parameters t_i. The controllability on Lie groups from a mathematical point of view was studied in [6,7,8,9]. This direct relation between the problem of controllability in Lie groups and the problem of universal quantum computation allows us to approach quantum computation with an alternative way beyond the "standard circuit model". In this approach if the selected Hamiltonians $J_1,\ J_2,\ \ldots, J_m$ form a complete set of operators, then every $W \in SU(2^N)$ can be exactly realized using a finite number of steps, although this number of steps is not fixed, where in the case of the "standard circuit model" the same element $SU(2^N)$ could be approximately realized using an infinitely number of steps. The order of generation (the number of computational steps required to perform an arbitrary N-qubit gate) is available for arbitrary Hamiltonians only in the case of the $SU(2)$ group (one-qubit gates) via the Lowental's criterion [10]. In this case, only two Hamiltonians $\{J_1,\ J_2\}$ are sufficient to form a complete set. If the Hamiltonians are orthogonal, i.e. $\mathrm{Trace}(J_1 J_2) = 0$, then three at most steps are required, to realizing any $W \in SU(2)$). When $\mathrm{Trace}(J_1 J_2) \neq 0$, the number of steps are given by the Lowental's criterion, but the algorithm to obtain the solution is not known.

In the case of higher order groups there is an analysis based on the Cartan decomposition of the $su(2^N)$ algebras [8]. This analysis provides also an analytical way of calculating the corresponding time parameters (Euler angles) in the case of the $SU(4)$ (2-qubit gates). On the same spirit is the proposal for exact computation by Whaley and collaborators [11]. Open problems in Lie groups controllability are:

a) *SU(2) group (one-qubit gates).* An algorithm which, given an arbitrary couple of generators–Hamiltonians, will be able to provide analytically the time parameters to perform universal computation, if the number of required steps are more than three.
b) *Higher order groups*. A criterion for minimum number of steps to generate an arbitrary element of the group (which corresponds to an arbitrary N-qubit gate, respectively) in the case where the generators–Hamiltonians are not orthogonal. Algorithms to evaluate the corresponding time parameters.

3. Quantum Gates Using the Intrinsic Abilities of a Physical System

The main points of our proposal can be summarized briefly as follows

- Instead of forcing a physical system to act as a predetermined set of universal gates we focus on the ability of the physical system to act as a quantum computer using only its natural available interactions (encoded universality [12]).
- Construction of any given gate and algorithm in terms of a minimal configuration and computational procedure.
- Minimized finite number of steps, evolving in time according to a finite number of basic, intrinsic Hamiltonians, controlled by a minimal finite number of classical switches (the selection of the switches is not unique).
- Implementation does not depend on the psecific system used as a qubit. Several solid state proposals as charge Josephson junctions, SQUIDs, quantum dots have been tested but our proposal can be extended to NMR quantum computation, trapped ions etc in order to test it with various physical systems described by different Hamiltonians and interactions.

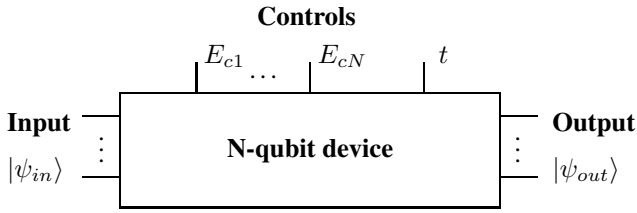


Figure 1. *Abstract N-qubit device*

This computer consists of one cell controlled by external binary switches and evolving in time using these switches. Quantum gates and algorithms are translated into manipulation of these switches. It is a simple device which overcomes the notion of transposition from classical circuits and does not have any "quantum" connections (one of the difficult parts in physical implementation-especially in solid state devices).

The above proposal is based in the following mathematical Propositions:

Proposition 3.1 *A number of $N+1$ switches are sufficient for universal quantum computation in a N-qubit device.*

Proposition 3.2 *A set of $N+2$ Hamiltonians provided by the basic Hamiltonian of the N-qubit device through appropriate tuning of the $N+1$ switches, can generate the $su(2^N)$ algebra.*

Proposition 3.3 *The construction scheme of any quantum gate consists of a finite number of steps evolving in time according to a finite number of basic Hamiltonians (reiterating in cyclic pattern) and provided by proper switches' manipulation.*

$$U = \ldots e^{-it_4 H_4} e^{-it_3 H_2} e^{-it_2 H_3} e^{-it_1 H_1} \tag{2}$$

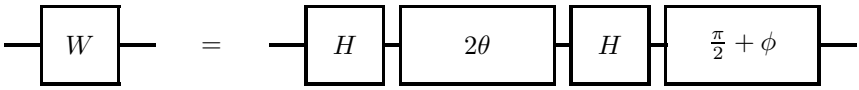
It is an open conjecture that every U can be generated exactly by $\mathcal{O}(4^N)$ steps.

The manipulations of the quantum computer can be codified by a rudimentary *Quantum Machine Language* [13].

4. Results

4.1. One-Qubit Gates

According to the "standard model" the most usual analysis of an arbitrary one-qubit gate includes two Hadamard gates and two phase gates.



Even if both Hadamard and phase gate can be realized in one computational step, then at least four steps are required to perform universal quantum computation. Usually decompositions of elementary one-qubit gates require at least 8 computational steps according to this analysis and it is performed most of the times in systems which provide us with orthogonal Hamiltonians.

With the presented computational scheme the results in the case of the one-qubit gates are the following:
Orthogonal Hamiltonians If the Hamiltonians are orthogonal, i.e. $(H_1, H_2) = \mathrm{Trace}(H_1 H_2) = 0$, then two Hamiltonians and three computational steps at most are required, to realize any $W \in SU(2)$.

$$W = e^{-it_3 H_1} e^{-it_2 H_2} e^{-it_1 H_1} \tag{3}$$

For example in the case of NMR (where orthogonal Hamiltonians are used) we can perform universal quantum computation within three computational steps, while the analytical solutions for the time parameters are the trivial Euler angles.

Non-Orthogonal Hamiltonians If the Hamiltonians are non-orthogonal i.e. $(H_1, H_2) = \mathrm{Trace}(H_1 H_2) \neq 0$, the number of steps–the order of generation is $n = k + 2$, given by the Lowental's criterion

$$\cos(\frac{\pi}{k}) < \frac{|(H_1, H_2)|}{(H_1, H_1)^{1/2}(H_2, H_2)^{1/2}} \leq \cos(\frac{\pi}{k+1}), \quad k \geq 2 \tag{4}$$

and the corresponding construction scheme is the following

$$W = e^{-it_n H_1} \ldots e^{-it_3 H_1} e^{-it_2 H_2} e^{-it_1 H_1} \tag{5}$$

For example in the case of the charge Josephson junctions where the general Hamiltonian is $H = \frac{1}{2} E_c \sigma_z - \frac{1}{2} E_J \sigma_x$ manipulation of the bias energy E_c which is controlled by the binary switch of gate voltage V_g, provides the following non-orthogonal Hamiltonians

$$H_1 = -\frac{1}{2} E_J \sigma_x \quad \text{and} \quad H_2 = \frac{1}{2} E_c \sigma_z - \frac{1}{2} E_J \sigma_x \tag{6}$$

The pair $\{H_1, H_2\}$ generates the $su(2)$ algebra but since $\mathrm{Trace}(H_1 H_2) \neq 0$ the whole $SU(2)$ group cannot be cover in 3 steps.

The Lowenthal's parameter ψ is

$$\psi = \frac{|(H_1, H_2)|}{(H_1, H_1)^{1/2}(H_2, H_2)^{1/2}} = \frac{\frac{E_J}{E_c}}{\sqrt{1+\frac{E_J^2}{E_c^2}}} = \frac{x}{\sqrt{1+x^2}} \quad \text{where } x = \frac{E_J}{E_c} \tag{7}$$

If x is small enough, then $\psi < \cos\frac{\pi}{3}$ and every element of the $SU(2)$ group $W = w_0\mathbb{I} - i(w_1\sigma_x + w_2\sigma_y + w_3\sigma_z)$ (one-qubit gate), can be constructed in 4 steps at most. The corresponding analytical solutions in that case are

$$t_1 = -\frac{2}{E_J}\arctan\left(\frac{Ec(w_0w_3-w_1w_2)+\sqrt{w_2^2+w_3^2}\sqrt{E_c^2(w_0^2+w_1^2)-E_J^2(w_2^2+w_3^2)}}{-E_c(w_0w_2+w_1w_3)+E_J(w_2^2+w_3^2)}\right)+\frac{4k_1\pi}{E_J}$$

$$t_2 = -\frac{2}{\sqrt{E_c^2+E_J^2}}\arctan\left(\frac{\sqrt{E_c^2+E_J^2}\sqrt{w_2^2+w_3^2}}{\sqrt{E_c^2(w_0^2+w_1^2)-E_J^2(w_2^2+w_3^2)}}\right)+\frac{4k_2\pi}{\sqrt{E_c^2+E_J^2}}$$

$$t_3 = -\frac{2}{E_J}\arctan\left(\frac{Ec(w_0w_3+w_1w_2)+\sqrt{w_2^2+w_3^2}\sqrt{E_c^2(w_0^2+w_1^2)-E_J^2(w_2^2+w_3^2)}}{E_c(w_0w_2-w_1w_3)+E_J(w_2^2+w_3^2)}\right)+\frac{4k_3\pi}{E_J}$$

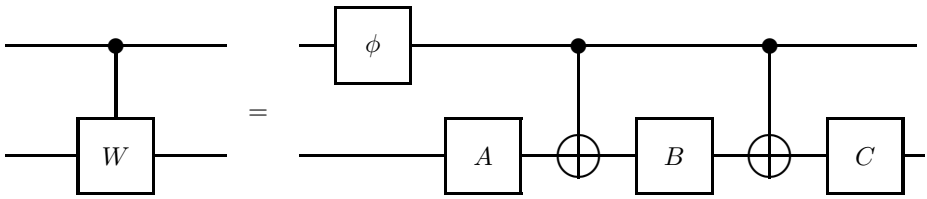
$$t_4 = \frac{2}{\sqrt{E_c^2+E_J^2}}\operatorname{arccot}\left(\frac{-2(\mathrm{w_1w_2+w_0w_3})\sqrt{1+\mathrm{x^2}}}{2(\mathrm{w_0^2+w_1^2-(w_2^2+w_3^2)\,x^2})}+\right.$$

$$\left.+\frac{\sqrt{4(w_1w_2+w_0w_3)^2(1+x^2)-4(w_0^2+w_1^2-(w_2^2+w_3^2)\,x^2)(2(w_0w_2-w_1w_3)\,x-(w_2^2+w_3^2)(-1+x^2))}}{2(w_0^2+w_1^2-(w_2^2+w_3^2)\,x^2)}\right)+$$

$$+\frac{4k_4\pi}{\sqrt{E_c^2+E_J^2}} \quad \text{where } k_1, k_2, k_3, k_4 \in \mathbb{N}$$

4.2. Two Qubit Gates

Analysis according to the "standard circuit model" requires at least 5 Hamiltonians and 19 computational steps and and it is performed most of the times in systems which provides orthogonal Hamiltonians.



In the case of orthogonal Hamiltonians there is the Cartan decomposition of the $SU(2^N)$ group [8], directly applied to the $SU(4)$ group, which gives analytical solutions and was recently extended with an algorithm to realize every $SU(2^N)$ [14]. The decomposition provided by [8] requires 5 different Hamiltonians and 27 computational steps to simulate an arbitrary gate while the number of computational steps reduces to 19 in the case of a controlled gate.

Next we show the results of numerical simulations of the present computational scheme:

Orthogonal Hamiltonians If the Hamiltonians are orthogonal (e.g. Heisenberg interaction [17], BQHD [18], SQUIDs [15]) then with two binary switches providing us with

3 different Hamiltonians and within 15 computational steps we cover the $SU(4)$ group (conjecture) and all the tested gates are successfully simulated.

For example, in a system described by a general Hamiltonian of the form $H = \sum_i^N \bar{B}^i(t)\hat{\sigma}^{(i)} + \sum_{i\neq j} J_{ab}^{ij}(t)\hat{\sigma}_a^{(i)}\hat{\sigma}_b^{(j)}$ (Heisenberg interaction), only 3 Hamiltonians

$$H_1 = B^1\,\sigma_z^{(1)}$$

$$H_2 = B^2\,\sigma_x^{(2)}$$

$$H_3 = J_{12}\,\left(\sigma_x^{(1)}\sigma_x^{(2)} + \sigma_y^{(1)}\sigma_y^{(2)} + \sigma_z^{(1)}\sigma_z^{(2)}\right) \tag{8}$$

are sufficient for universal quantum computation in 15 computational steps

$$U = e^{H_3 t_{15}} e^{H_2 t_{14}} e^{H_1 t_{13}} \ldots e^{H_3 t_3} e^{H_2 t_2} e^{H_1 t_1} \tag{9}$$

Non-Orthogonal Hamiltonians If the Hamiltonians are non-orthogonal (charge Josephson junctions [15], quantum dots [16], permanent interaction which cannot be switched off etc) but the interaction between the qubits is weak, then using 4 different Hamiltonians and within 15 computational steps (time parameters) a large part of the $SU(4)$ is covered and all the known important gates for quantum computation are successfully simulated. In general, the weaker the interaction, the larger the part of the group covered (more gates can be simulated).

a) *Permanent Interaction*. If the interaction J_{12} of the previous paradigm can not be switched of then a construction scheme with two binary switches and 3 non-orthogonal Hamiltonians

$$H_1 = B^1\,\sigma_z^{(1)} + J_{12}\,\left(\sigma_x^{(1)}\sigma_x^{(2)} + \sigma_y^{(1)}\sigma_y^{(2)} + \sigma_z^{(1)}\sigma_z^{(2)}\right)$$

$$H_2 = B^2\,\sigma_x^{(2)} + J_{12}\,\left(\sigma_x^{(1)}\sigma_x^{(2)} + \sigma_y^{(1)}\sigma_y^{(2)} + \sigma_z^{(1)}\sigma_z^{(2)}\right)$$

$$H_3 = J_{12}\,\left(\sigma_x^{(1)}\sigma_x^{(2)} + \sigma_y^{(1)}\sigma_y^{(2)} + \sigma_z^{(1)}\sigma_z^{(2)}\right) \tag{10}$$

simulates all the basic gates in 15 steps (9).

b) *Charge Josephson junctions*. A system of two identical coupled Josephson junctions is described by the following general Hamiltonian $\frac{1}{2}E_{c_1}\,\sigma_z^{(1)} - \frac{1}{2}E_{J_1}\,\sigma_x^{(1)} + \frac{1}{2}E_{c_2}\,\sigma_z^{(2)} - \frac{1}{2}E_{J_2}\,\sigma_x^{(2)} - \frac{1}{2}E_L\,\sigma_y^{(1)}\,\sigma_y^{(2)}$. Manipulation of 3 binary switches of the system provides the following 4 non-orthogonal Hamiltonians

$$H_1 = \frac{1}{2}E_c\,(\sigma_z^{(1)} + \sigma_z^{(2)}) - \frac{1}{2}E_J\,(\sigma_x^{(1)} + \sigma_x^{(2)})$$

$$H_2 = -\frac{1}{2}E_J\,(\sigma_x^{(1)} + \sigma_x^{(2)}) - \frac{1}{2}E_L\,\sigma_y^{(1)}\sigma_y^{(2)}$$

$$H_3 = \frac{1}{2}E_c\,\sigma_z^{(2)} - \frac{1}{2}E_J\,(\sigma_x^{(1)} + \sigma_x^{(2)})$$

$$H_4 = \frac{1}{2}E_c\,\sigma_z^{(1)} - \frac{1}{2}E_J\,(\sigma_x^{(1)} + \sigma_x^{(2)}) \tag{11}$$

and all basic gates are simulated in 15 steps (9).

The efficiency of our simulation is defined by a test function, f_{test}. It is a function of 15 time variables

$$f_{test}(t_1,\ldots,t_{15}) = \sum_{i,j=1}^{4} |(U_{\rm gate})_{ij} - (U(t_1,\ldots,t_{15}))_{ij}|^2 = ||U_{\rm gate} - U||^2 \tag{12}$$

In our numerical simulations f_{test} usually attains values of 10^{-8} or less. Taking into account more decimal digits in the approximation of the time parameters results to a further decrease of its value. Gates that have been tested numerically are all the important two-qubit gates for quantum computation such as the CNOT gate, the SWAP gate, the Quantum Fourier Transform gate for two qubits, several controlled gates etc. The ratio of the values of the external switches tuning amplitudes over the magnitude of the interaction is not in the area of hard pulses. Numerical results are available upon request from the authors.

Acknowledgements

The work of K.Ch. Ch. was supported by Herakleitos Research Scholarships (21866). The authors would like to thank I. Jex and G. Chadjitaskos of the Czech Technical University in Prague, D. Angelakis and S. G. Schirmer of the Centre for Quantum Computation (DAMTP-University of Cambridge) for fruitful discussions.

References

[1] A. Barenco et al, *Phys. Rev. A* **52**, (1995) 3457–3467.
[2] D. Deutsch, *Proc. Roy. Soc. London A* **400**, (1985) 97–117.
[3] S. Lloyd, *Phys. Rev. Lett.* **75**, (1995) 346–349.
[4] D. Deutsch et al, *Proc. R. Soc. London A* **449**, (1995) 678.
[5] A. Spröl et al, e-print quant-ph/0504202.
[6] V. Jurdjević, H. Sussmann, *J. Diff. Eq.* **12**, (1975) 313-329.
[7] V. Ramakrishna et al, *Phys. Rev. A* **51**, (1995) 960–966.
[8] N. Khaneja et al, *Chem. Phys.* **267**, (2001) 11–23.
[9] S. G. Schirmer et al, *Phys. Rev. A* **63**, (2001) 025403.
[10] F. Lowenthal, *Can. J. Math.* **24**, No 4, (1972) 713–727.
[11] J. Zhang et al, *Phys. Rev. Lett.* **91**, (2003) 027903.
[12] D. P. DiVincenzo et al, *Nature* **408**, (2000) 339–342.
[13] K. Ch. Chatzisavvas et al, *Quantum Computers and Computing*, **4**, No2, (2003) 93–110.
[14] H. N. Sá Earp et al, *J. Math. Phys.* **46**, (2005) 082108.
[15] Y. Makhlin et al, *Rev. Mod. Phys.* **73**, (2000) 357–400.
[16] T. Tanamoto, *Phys. Rev. A* **61**, (2000) 22305.
[17] S. Braunstein and H. K. Lo, *Fort. Phys.* **48**, (2000) 765–1138.
[18] V. W. Scarola et al, *Phys. Rev. Lett.* **91**, (2003) 167903.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Entangled Light From Optical Time Boundaries

A. Guerreiro [a,1], A. Ferreira [a] J. T. Mendonça [b] and V. Vedral [c]
[a] *Dep. Física Fac. Ciências Univ. Porto, Campo Alegre 687, 4169-007, Porto, Portugal*
[b] *Inst. Sup. Técnico, Av. Rovisco Pais 1, 1000-69, Lisboa, Portugal*
[c] *The School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT, UK*

Abstract. We discuss a mechanism of generating two separable beams of light entangled in momentum using a sharp optical temporal boundary to excite Zero Point Fluctuations from the vacuum state of the electromagnetic field. This phenomenon is known as Time Refraction. Similarities between this effect and the Unruh effect are discussed.

Keywords. Entanglement, Quantum fields, Time Refraction, Unruh effect

1. Introduction

Throughout the past half century, work on topics such as the Unruh radiation [1] in accelerated reference frames and Hawking evaporation of a black hole [2] has explored how Zero Point Fluctuations (ZPF) can be excited and directly observed. In particular, the Unruh effect[1] predicts that an observer in an accelerating frame detects thermal radiation corresponding to the vacuum state of an inertial observer. Although initially treated as a purely kinematic effect, the corresponding dynamical effect with actual emission of radiation was treated in later papers [3], which led to the proposal of several experimental schemes [4][5]. Another possibility is the dynamical Casimir effect, which predicts the production of photons by resonantly amplifying ZPF in an optical cavity with moving end mirrors [6]. Though the Unruh radiation and the dynamical Casimir effect are fundamental and macroscopic manifestations of ZPF and they are well understood theoretically, it has not yet been possible to measure these extremely weak processes. For example, it would be necessary to create acceleration gradients of $10^{20} m/s^2$ to obtain Unruh radiation corresponding to a temperature of $1K$.

Recently, a series of papers [7] have introduced the concept of Time Refraction, which describes how the classical and quantum properties of light are altered by the sudden change of the optical properties of a medium. Time Refraction results from the violation of the symmetry of time inversion and extends the usual concept of refraction into the time domain, while Unruh and Hawking radiations explore the lack of equivalence between different types of observers. Like the Unruh effect [1], the Hawking mechanism

[1]Correspondence to: A. Guerreiro, Departamento de Física da Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal; Tel.: +359 22 608 2611; Fax: +359 608 2628; E-mail:ariel@fc.up.pt

[2] and the dynamical Casimir effect [6], the quantum theory of Time Refraction predicts the excitation of virtual particles from the turmoil of ZPF and the emission of real particles.

In a very recent work [8], a different process of excitation of ZPF was proposed based on a non-accelerated optical boundary moving with *apparent* superluminal velocity across an optical medium, which also leads to the emission of photons, but now the optical boundary works as a relativistic partial mirror producing a Doppler shift and an increase in the number of photons emitted. This paper addresses the generation of entanglement in the process of Time Refraction.

2. Time Refraction

Let us consider an infinite, nondispersive, and nondissipative dielectric medium, characterized by the refractive index n_1, which at time $t = 0$ is perturbed, resulting in a new value of the refractive index n_2. This can be described by a time-dependent refractive index $n(t) = n_1 H(-t) + n_2 H(t)$, where $H(t)$ is the Heaviside function. The field modes for $t < 0$ and $t > 0$ are different since the electromagnetic field must satisfy different dispersion relations. Also, and from a classical point of view, for the dielectric displacement and the magnetic induction field to satisfy the macroscopic Maxwell equations they must be continuous at time $t = 0$ [9][7]. This implies that the time boundary can be described as a four port device, coupling two initial complex plane wave modes: $\phi_i(\overrightarrow{r}) = \exp[-i\overrightarrow{k}_i.\overrightarrow{r}]$ and $\phi_a(\overrightarrow{r}) = \exp[-i\overrightarrow{k}_a.\overrightarrow{r}]$ existing for $t < 0$, with two final complex plane wave modes $\phi_t(\overrightarrow{r}) = \exp[-i\overrightarrow{k}_t.\overrightarrow{r}]$ and $\phi_r(\overrightarrow{r}) = \exp[-i\overrightarrow{k}_r.\overrightarrow{r}]$ existing for $t > 0$, which satisfy $\overrightarrow{k}_t = (n_2/n_1)\overrightarrow{k}_i$, $\overrightarrow{k}_r = (n_2/n_1)\overrightarrow{k}_i = -\overrightarrow{k}_t$, $\overrightarrow{k}_a = -\overrightarrow{k_i}$. From the quantum point of view, this implies that the annihilation and creation operators for these modes can be related as:

$$a_i = Aa_t - Ba_r^+ \text{ and } a_i^+ = Aa_t^+ - Ba_r, \tag{1}$$

$$a_a = Aa_r - Ba_t^+ \text{and } a_a^+ = Aa_r^+ - Ba_t, \tag{2}$$

where $A = \left(1+\alpha^2\right)/2\alpha$, $B = \left(1-\alpha^2\right)/2\alpha$, $\alpha = (n_1/n_2)^{1/2}$, satisfying $A^2 - B^2 = 1$. These relations describe a squeezing transformation $S = \exp\left[r\left(a_i a_a - a_i^+ a_a^+\right)\right]$, with parameter $r = \cosh^{-1}(A) = \sinh^{-1}(B)$ and they are formally identical to the Bogolyubov transformation [14] between the field modes in the inertial and accelerated frames associated with an Unruh temperature $T = \hbar w/(K_B \ln[\left(1-\alpha^2\right)/\left(1+\alpha^2\right)])$, where K_B is the Boltzmann constant and w is the photon frequency.

3. Vacuum Radiation and Entanglement

As demonstrated in references [10][11], transformations (1) and (2) imply that, after the optical perturbation has passed, the initial vacuum state can be expressed in terms of the new eigenstates of the field according to $|0\rangle_i |0\rangle_a = \sum_n C_n |n\rangle_t |n\rangle_r$with $C_n = (1-|z|)^{1/2} z^{n/2}$ and $z = (B/A)^2 = \tanh^2(r)$. This implies the emission of pairs of photons moving along opposite directions with wave vectors $\overrightarrow{k}_t$ and $\overrightarrow{k}_r = -\overrightarrow{k}_t$ and with a mean number of photons,

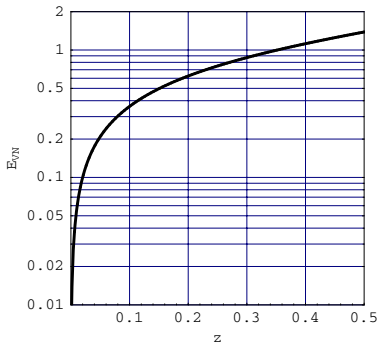


Figure 1. Entanglement entropy $E_{VN}(\rho_{S\otimes S'})$ as function of z.

$$\langle N_t\rangle = \langle N_r\rangle = z/(z-1). \tag{3}$$

Equations (1) and (2) change the vacuum state into another pure state for which the Entanglement Entropy is a suitable entanglement measure [12]. The Entanglement Entropy $E_{VN}(\rho_{S\otimes S'})$ of a system composed by two sub-systems S and S' and described by a density operator $\rho_{S\otimes S'}$ is calculated by tracing out one of the sub-systems in the density operator, i.e. $\rho^S = Tr_{S'}\left[\rho_{S\otimes S'}\right]$ and calculating $E_{VN}(\rho_{S\otimes S'}) = -K_B\, Tr\left[\rho^S \ln(\rho^S)\right]$. Here the subsystems are to the two distinct sets of photons emitted, corresponding to ϕ_r and ϕ_t, and E_{VN} yields:

$$\begin{aligned} E_{VN} &= -K_B \ln(1-z) - K_B z \ln(z)/(1-z) \\ &= K_B(\langle N_t\rangle + 1)\ln(\langle N_t\rangle + 1) - K_B \langle N_t\rangle \ln[\langle N_t\rangle]. \end{aligned} \tag{4}$$

Notice that the E_{VN} is basically the Shannon entropy introduced by increasing one photon pair into the system and, naturally, is proportional to the number of photons produced. A plot of the Entanglement Entropy is presented in Figure 1. The entanglement increases monotonically with z, i.e. with the amount of change of the refractive index.

4. Conclusion

In conclusion, we have presented a new mechanism for the emission of entangled radiation (entanglement in momentum) using a sharp optical perturbation in a medium. The photon emission and entanglement produced by Time Refraction [7] is limited by the maximum variation of refractive index produced by an optical perturbation, which usually is of the order of less than 1%, yielding a very weak effect. However, other similar effects, namely the emission of light from optical boundaries with apparent superluminal velocities (which could be called Superluminal Space-Time Refraction) [8] and Temporal Intereference [13] could be used to enhance these effects into measurable experimental parameters.

Finally, it should be noted that a dielectric medium is optically equivalent to a gravitational field [10]. The creation of pairs of photons in vacuum by the mechanism of time refraction is connected with the Unruh effect [1] and with the Hawking mechanism of black-hole evaporation [2]. In fact, equations (1) and (2) are formally identical to the Bouglioubov transformation between the field modes in the inertial and accelerated frames derived in Unruh effect [1][14]. The main difference being that, while for Time Refraction we can have access to both the time-transmitted and time-reflected waves, an accelerated observer can only measure a reduced set of modes due to the existence of an horizon, resulting in an intrinsic entropy which may destroy the entanglement. From the fundamental point of view, both these radiative process are identical and they result from the condensation of an infinite number of virtual photons of the initial vacuum into a finite number of real photons emitted from a new vacuum state. These types of process are well understood by thermal field dynamical theory [15] and are commonly associated with spontaneous symmetry breaking.

The resemblance between Time Refraction and the Unruh effect raises intriguing questions regarding the relation between the Unruh effect and entanglement. In particular, we could ask under which circumstances the generation of entanglement via Time Refraction can be transposed into the case of the Unruh effect and what the physical meaning of the generation of entanglement from ZPF is.

Acknowledgements

We gratefully acknowledge stimulating discussions with Č. Brukner, C. Herdeiro, J. Kofler and M. Wieśniak. A.F. is supported by FCT (Portugal) through grant PRAXIS no. SFRH/BD/18092/04. V.V. acknowledges funding from EPSRC and the European Union.

References

[1] W.G. Unruh, *Phys. Rev. D*, **14**, 870 (1976).
[2] S.W. Hawking, *Nature(London)*, **248** 30 (1974), *Commn. Math Phys*, **43** 199 (1975).
[3] P. Chen and T. Tajima, *Phys. Rev. Lett.*, **83** , 256 (1999).
[4] E. Yablonovich, *Phys. Rev. Lett.*, **62** , 1742 (1989), *Phys. Rev. A*, **10**, 1888 (1974).
[5] S. M. Darbinyan, K. A. Ispiryan, M.K. Ispiryan and A.T. Margaryan, *JETP*, **51**, 110 (1990).
[6] V. V. Dodonov, A. B. Klimov, V. I. Man'ko, *Phys. Lett. A*, **149**, 225 (1990); V. V. Dodonov, *Phys. Rev. Lett.*, **207**, 126 (1995); R. Schültzhold, G. Plunien, and G. Soff, *Phys. Rev. A*, **57**, 2311 (1998).
[7] J. T. Mendonça, A. Guerreiro and A.M. Martins, *Phys. Rev. A*, **62**, 033805 (2000); J.T. Mendonça, *The Theory of Photon Acceleration.*,Institute of Physics Publishing Bristol (2000).
[8] A. Guerreiro, J. T. Mendonça and A. M. Martins, *J. Opt. B: Quantum Semiclass. Opt.*, **7**, S69(2005).
[9] M. Cirone, K. Rgazewski, and J. Mostowski, *Phys. Rev. A,* **55**, 62 (1997).
[10] U. Leonhardt and P. Piwnicki, *Phys. Rev. Lett.,* **84** , 822 (2000).
[11] L. Gilles and P. L. Knight, *J. Mod. Opt*, **39** , 1411 (1992).
[12] C. H. Bennett, H. J. Bernstein, S. Popescu and B. Schumacher, *Phys. Rev. A* **53**, 2046 (1996).
[13] J. T. Mendonça, A. M. Martins and A.Guerreiro, *Phys. Rev. A* **68** 043801 (2003).
[14] P. M. Alsing, G. J. Milburn, *Phys. Rev. Lett.* **91**, 180404 (2003).
[15] H. Umezawa, *Advanced Field Theory.*, American Institute of Physics(1992).

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Strong Light-Matter Coupling: Parametric Interactions in a Cavity and Free-Space

I. B. Mekhov [a,1], V. S. Egorov [a], V. N. Lebedev [a], P. V. Moroshkin [a], I. A. Chekhonin [a], S. N. Bagayev [b]

[a] *St. Petersburg State University, V.A. Fock Institute of Physics, Russia*
[b] *Institute of Laser Physics, Siberian Branch of RAS, Novosibirsk, Russia*

Abstract. We consider parametric interactions of laser pulses in a coherent macroscopic ensemble of resonant atoms, which are possible in the strong coupling regime of light-matter interaction. The spectrum condensation (lasing at collective vacuum Rabi sidebands) was studied in an active cavity configuration. Parametric interactions under the strong light-matter coupling were proved even in free space. In contrast to bichromatic beats in a cavity, they were shown to appear due to interference between polaritonic wave packets of different group velocities.

Keywords. Quantum optics, dense atomic ensembles, Dicke superradiance, collective effects, resonant parametric processes, polaritons

1. Introduction

We present a study of parametric interactions between laser pulses in a coherent ensemble of two-level quantum objects ("atoms"). We concentrate on the processes that can be obtained in the strong coupling regime of the light-matter interaction, which is achieved, when (i) the frequency of the field-matter excitation exchanges exceeds the decoherence rates, (ii) the external field is not strong enough to entirely determine the evolution of a system: it is the reemission field that plays a key role and provides the collective behaviour of atoms.

Recently, this regime has attracted attention in quantum optics with both atoms and nanostructures such as quantum wells and dots (cavity QED, Dicke effects, microcavity exciton-polariton parametric scattering, squeezing and entanglement [1]). In both fields, it is considered as one of the key models for quantum information processing (QIP). In this context, the works on single or small numbers of objects (photons, atoms, excitons) as qubits should be mentioned [2,3].

During the last few years, there has been a growing interest in the study of macroscopic ensembles of quantum objects, which can be used as collective elements for QIP protocols [4,5]. Quantum memories [4,6,7], sources of single photons [8], entanglement of ensembles [9] and new sources of entangled fields based

[1] Correspondence to: Igor Mekhov, E-mail: Mekhov@yahoo.com.

on optically dense atomic ensembles were recently proposed. The advantages of such collective objects in contrast to single-particle qubits were highlighted [10].

Investigations of collective effects in the strong coupling regime may be of particular importance because in this regime some purely quantum properties of the phenomena do not decrease with increasing number of particles [11,12], which will be important for QIP with both discrete and continuous variables. In Ref. [1], a scheme for generation of twin polaritons (coupled light-matter excitations) in semiconductor microcavities with quantum wells was developed. It was based on the parametric scattering of microcavity exciton-polaritons in the strong coupling regime. In this scheme, very strong quantum correlations between polaritons were obtained inside a cavity. Nevertheless, since only one polariton branch of the dispersion curve was used in the parametric process proposed, quantum correlations in the outgoing light fields were essentially reduced. In this report, we present a study of parametric processes both in a cavity and free space, which arise due to collective energy exchange between field and matter systems (i.e., beating between two polariton branches of the dispersion curve) [13,14,15].

2. Interactions in a Cavity

The most significant manifestation of the strong coupling regime consists of the resonant density-dependent splitting of the dispersion curve into two polariton branches. The splitting appears when the collective coupling constant (cooperative frequency) of a medium ω_c exceeds all rates of incoherent relaxations γ:

$$\omega_c = \sqrt{2\pi d^2 \omega_0 n/\hbar} = g\sqrt{N} \gg \gamma, \tag{1}$$

where d and ω_0 are the dipole moment and frequency of a transition, n and N are the density and number of atoms and g is the single-atom coupling constant. Thus, in dense media, photons with equal wave vectors but different frequencies exist.

If a spatial spectrum of a problem is fixed by a single-mode cavity, beating between these frequencies corresponds to vacuum Rabi oscillations. A spectral doublet, arising in this case, represents density-dependent vacuum Rabi splitting of a cavity mode, which is proportional to ω_c (1). The origin of this effect can be traced to the linear interaction of light and dipoles in a passive cavity [12]. The weak field condition (which cannot destroy polariton dispersion) is reduced to the statement that the photon number is smaller then the number of atoms.

In this report, we present results related to the strong light-matter coupling in an active system: a multimode broadband laser with an intracavity narrowband coherent resonant medium without population inversion. The condition of strong coupling (1) can be fulfilled with respect to that macroscopic resonant absorber. A phenomenon of spectrum condensation was analysed. It consists of the dramatic modification of a lasing spectrum under transition from the weak- to strong-coupling regime of the light-matter interaction: instead of the usual saturated absorption line, a bright narrow doublet of generation appears.

Figure 1 presents dye laser spectra obtained with a neon discharge as an intracavity absorber at different densities n (up to 10^{13} cm^{-3} corresponding to

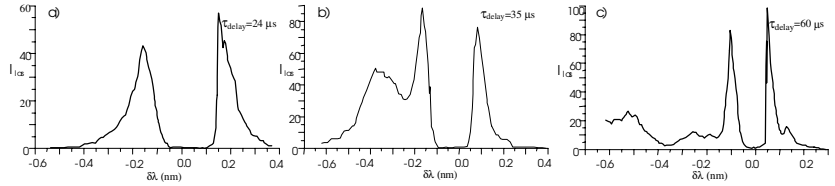


Figure 1. Self-splitting of a dye laser spectrum at different moments of a Ne discharge afterglow, which corresponds to decreasing atomic density in an intracavity cell, $\lambda = 594.5$ nm.

$\omega_c/2\pi$ up to 10 GHz) of neon atoms in the metastable state, which was used as a lower state of a two-level system. It was shown that the density dependence of the doublet splitting is in agreement with that of ω_c (1), which determines the vacuum Rabi splitting. Moreover, the the existence of a density threshold, which is also determined by Eq. (1), was demonstrated. On the basis of the cooperative parametric resonance model [13], this effect was treated as lasing at collective vacuum Rabi sidebands due to parametric interactions between laser modes.

3. Strong Coupling Regime in Free Space

The main difference of the free-space interaction results from the fact that the spatial spectrum of a problem is not fixed by a cavity. As a consequence, under the linear propagation of a short pulse, no coherent density-dependent spectral features (such as vacuum Rabi splitting in a cavity) can be extracted from spectral measurements, except for the trivial appearance of an incoherent absorption line. Nevertheless, in our work we have shown that such coherent features can be obtained under nonlinear parametric interaction of laser pulses (the details of the theory can be found in Ref. [15], experimental results are presented in Ref. [14]). In contrast to cavity oscillations, they were shown to appear due to free-space optical ringing, which does not originate from beating between waves of equal wave vectors, but from the successive beats between polaritonic wave packets of equal group velocities. Long coherent oscillations can be obtained due to a considerable reduction of the group velocity in a dense medium and due to its essential variation over the broad spectrum of a short laser pulse. The oscillation frequency at the initial stage of the ringing ω_D should exceed the decoherence rates γ:

$$\omega_D = \omega_c^2 z/c \gg \gamma. \tag{2}$$

This frequency increases with propagation distance z and medium density n. Thus, it is proportional to the number of interacting atoms, which directly demonstrates the superradiant character of the field reemission by the atomic ensemble.

We considered the nonlinear parametric interaction between two short laser pulses propagating in a dense resonant atomic ensemble. Optical ringing was shown to significantly affect the propagation and amplification of a probe field under its interaction with a nearly copropagating pump. Depending on the probe-pump time delay τ_0 (the probe pulse precedes the pump for $\tau_0 < 0$), the probe

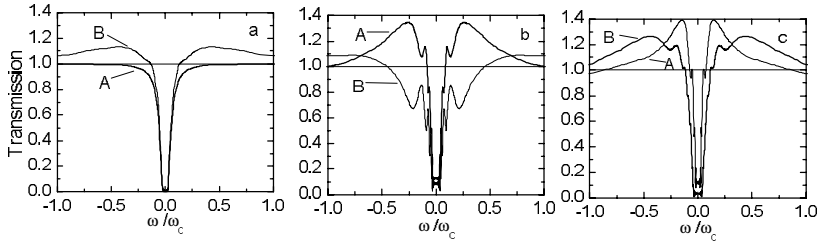


Figure 2. Transmission spectra of the probe field. (a) linear transmission (A) and $\omega_c\tau_0 = 0$ (B), $\omega_c z/c = 1.0$; (b) parametric amplification at $\omega_c\tau_0 = -0.5$ (A) and dumping at $\omega_c\tau_0 = 0.5$ (B) for $\omega_c z/c = 1.0$; (c) parametric amplification at $\omega_c\tau_0 = -0.5$ for $\omega_c z/c = 0.5$ (A) and for $\omega_c z/c = 2.0$ (B); $\gamma/\omega_c = 10^{-3}$, pump area $s_0 = 0.49\pi$.

transmission spectra show either a specific doublet or dip, which corresponds to parametric sideband amplification or dumping of radiation [cf. Fig. 2(b)]. The widths of these features are greater than the width of an incoherent absorbtion contour [shown in Fig. 2(a)], they are determined by the density-dependent field-matter coupling constant and increase during the propagation [cf. Figs. 2(b), (c)].

The spectral features in Fig. 2 can be explained as a result of parametric forward scattering of a probe pulse on spatiotemporal modulations of the population difference, appearing due to the optical ringing in a pump [15]. The characteristic frequencies, which increase with the number of interacting atoms, correspond to the frequency of optical ringing. The weak field condition, which does not destroy dispersion and the collective behaviour of atoms, in this case, is equivalent to the requirement that the input pulse area be small (so that 0π ringing should not be shadowed by 2π solitons and Rabi flopping). Contrary to strong-field parametric processes (e.g., due to transient Rabi flopping or stationary Mollow-Boyd effect, which are determined by the pump amplitude), the density- and coordinate-dependent spectra of the probe display the importance of free-space collective oscillations and cannot be obtained in the framework of a single-atom model.

References

[1] J. Ph. Karr, A. Baas, and E. Giacobino, Phys. Rev. A **69** (2004), 063807.
[2] J. M. Raimond, M. Brune, and S. Haroche, Rev. Mod. Phys. **73** (2001), 565.
[3] J. Vuckovic and Y. Yamamoto, Appl. Phys. Lett. **82** (2003), 2374.
[4] M.D. Lukin, Rev. Mod. Phys. **75** (2003), 457.
[5] L.-M. Duan, M.D. Lukin, J.I. Cirac, and P. Zoller, Nature **414** (2001), 413.
[6] D.V. Kupriyanov, O.S. Mishina, I.M. Sokolov, B. Julsgaard, E.S. Polzik, Phys. Rev. A **71** (2005), 032348; O.S. Mishina et al., report at this NATO ASI.
[7] K. Hammerer, K. Molmer, E.S. Polzik, J.I. Cirac, Phys. Rev. A **70** (2004), 044304.
[8] C.W. Chou, S.V. Polyakov, A. Kuzmich, and H.J. Kimble, Phys. Rev. Lett. **92** (2004), 213601.
[9] L.-M. Duan, J.I. Cirac, P. Zoller, and E.S. Polzik, Phys. Rev. Lett. **85** (2000), 5643.
[10] A. Kuzmich and T.A.B. Kennedy, Phys. Rev. Lett. **92** (2004), 030407.
[11] G. Rempe, R.J. Thompson, R.J. Brecha, W.D. Lee, and H.J. Kimble, Phys. Rev. Lett. **67** (1991), 1727.
[12] H.J. Carmichael, R.J. Brecha, P.R. Rice, Opt. Commun. **82** (1991), 73.

[13] V.V. Vasil'ev, V.S. Egorov, A.N. Fedorov, and I.A. Chekhonin, Opt. Spectrosc. **76** (1994), 146 [Opt. Spectrosc. **76** (1994), 134].
[14] S.N. Bagayev, V.S. Egorov, I.B. Mekhov, P.V. Moroshkin, I.A. Chekhonin, E.M. Davliatchine, and E. Kindel, Phys. Rev. A **68** (2003), 043812.
[15] V.S. Egorov, V.N. Lebedev, I.B. Mekhov, P.V. Moroshkin, I.A. Chekhonin, and S.N. Bagayev, Phys. Rev. A **69** (2004), 033804.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Macroscopic Quantum Information Channel Via the Polarization-Sensitive Interaction Between the Light and Spin Subsystems

Oksana S. Mishina [a,1], Dmitriy V. Kupriyanov [a] and Eugene S. Polzik [b]
[a] *Department of Theoretical Physics, State Technical University, St.-Petersburg, Russia*
[b] *QUANTOP - Danish Quantum Optics Center, Niels Bohr Institute, Copenhagen, Denmark*

Abstract. We discuss a quantum information channel which is based on a coherent and polarization sensitive interaction of light and atomic spin waves. We show that the joint Heisenberg dynamics of the Stokes polarization components of light and of the angular momenta of atoms has a wave nature and can be properly described in terms of the macroscopic polariton-type spin wave created in the sample. The principles of the quantum memory and readout protocols via the wave coupling of the output time modes in the light subsystem and the output spatial modes in the spin subsystem are demonstrated.

Keywords. Quantum fluctuations, Spin waves, Quantum memory

1. Introduction

Quantum information processing based on continuous variables as an alternative to the discrete quantum schemes has been intensely studied during the last decade, see, for example, the review [1] and references therein. Spin oriented atomic ensembles and polarized coherent light can be considered as convenient physical objects for storage and for transport of the quantum information, which can be written in the fluctuations of their polarizations. There are two experimental demonstrations of the entanglement and memory protocols in Refs. [2,3], where the dynamical coupling of the spin fluctuations of an atomic ensemble consisting of caesium atoms with the polarization fluctuations of linearly polarized light via the Faraday effect were implemented. The main feature and certain advantage of those protocols are that only collective integral variables of light and of atomic subsystems contributed in the quantum fluctuation interchange. However, as was recently reported in Ref. [4], in a more general situation the light and spin subsystems develop polariton-type spin wave dynamics in the sample. That means that the dy-

[1]Correspondence to: O.S. Mishina, Politehnicheskaya St. 29, 195251, Saint-Petersburg, Russia. Tel.: +7 812 552 6501; Fax: +7 812 247 2088; E-mail: m-sana@mail.ru.

namics of the Stokes polarization components and of the atomic spins becomes coupled not integrally and collectively but locally in time as well as in space during the whole interaction cycle. In this case the quantum correlations are spread among all the possible time and spatial polariton modes created in the sample. In this report we further develop the ideas of Ref.[4] and show how they could be applied for organizing the readout and memory protocols based on the wave dynamics of the process.

2. Macroscopic Polariton-Type Spin Wave

Consider an off-resonant pulse of radiation probing a spin-polarized atomic ensemble during the short time interval so that incoherent scattering is negligible. Then, as shown in [4], the relevant Heisenberg operators describing the light and atomic subsystems, and responsible for the quantum correlations dynamics, obey the following wave-type equations

$$\begin{aligned}
\left[\frac{\partial}{\partial z}+\frac{1}{c}\frac{\partial}{\partial t}\right]\hat{\Xi}_1(z,t) &= -\kappa_2\,\hat{\Xi}_2(z,t) \;+\; 2\beta\,\bar{\Xi}_3\,\hat{\mathcal{J}}_z(z,t)\\
\left[\frac{\partial}{\partial z}+\frac{1}{c}\frac{\partial}{\partial t}\right]\hat{\Xi}_2(z,t) &= \;\;\kappa_2\,\hat{\Xi}_1(z,t) \;-\; 2\epsilon\,\bar{\Xi}_3\,\hat{\mathcal{J}}_y(z,t)\\
\frac{\partial}{\partial t}\hat{\mathcal{J}}_z(z,t) &= \;\;\Omega\,\hat{\mathcal{J}}_y(z,t) \;-\; \epsilon\bar{\mathcal{J}}_x\,\hat{\Xi}_1(z,t)\\
\frac{\partial}{\partial t}\hat{\mathcal{J}}_y(z,t) &= -\Omega\,\hat{\mathcal{J}}_z(z,t) \;+\; \beta\bar{\mathcal{J}}_x\,\hat{\Xi}_2(z,t)
\end{aligned} \tag{1}$$

In these equations the light pulse with 100% linear polarization along x-direction propagates through the sample along z-direction. The atoms also have 100% orientation of their angular momenta along the x-direction. The field subsystem is described by the Heisenberg operators for two Stokes polarization components $\hat{\Xi}_1(z,t)$ and $\hat{\Xi}_2(z,t)$. The first Stokes component $\hat{\Xi}_1(z,t)$ is responsible for imbalance between photon fluxes of the modes linearly polarized along the ξ and η axes rotated with respect to x and y directions by an angle of $\pi/4$, and the second component $\hat{\Xi}_2(z,t)$ is responsible for imbalance in the right-hand and the left-hand polarizations. The third Stokes component $\bar{\Xi}_3$, responsible for imbalance in x- and y-type polarizations, is a quantum integral of motion and considered as external (non-operator) parameter in equations (1). The atomic subsystem is described by the operators $\hat{\mathcal{J}}_z(z,t)$ and $\hat{\mathcal{J}}_y(z,t)$, which are the Heisenberg operators for the spatial distribution of the transverse fluctuations of the angular momenta in respect to the z and y axes. The spin projection $\bar{\mathcal{J}}_x$ on x direction is another quantum integral of motion and is also considered a c-number parameter.

The system of equations (1) describes the coupled wave-type dynamics of the Stokes components of the light subsystem and of the transverse fluctuations of the atomic angular momenta. Since the field subsystem consists of many photons and the spin subsystem consists of macroscopic number of individual atomic spins, we shall call such a wave a macroscopic polariton-type spin wave. This polariton wave describes the mesoscopically smooth dynamics for the interacting microscopic quantum variables of the light and spin subsystems. Such a wave represents a manifestly quantum object, which exists only at the level of quantum description of the field and spin fluctuations. While deriving

equations (1) it was assumed that atoms preserve their location during the interaction cycle and there is no destruction of the coherent wave dynamics coming from the random atomic motion. This important condition can be fulfilled, for example, for interaction durations longer than a tenth of a millisecond for an ensemble consisting of ultra–cold atoms. Alternatively for hot atoms in a gas cell the interaction time should be shorter than a few microseconds.

The following two important parameters govern the dynamics of the process. The first one is β, which is the angle of the polarization rotation of the probe light due to the Faraday effect per one spin flip in the ensemble in z-direction. The second is ϵ, which is the ellipticity induced in the propagating light by the atomic sample due to the Cotton-Mouton effect per one spin flip in y-direction. The frequency $\Omega = \Omega_0 + \Omega_2$ combines the regular precession caused by the external magnetic field, Ω_0, with the frequency of light-induced shift of the Zeeman sublevels, Ω_2. The parameter κ_2 is responsible for birefringence effects with respect to x- and y-type polarizations of the probe, i.e. for the unitary transformation of linear polarization to circular polarization and vice versa.

In general the system of equations (1) can only be solved numerically. However, in the special case where at least one of the parameters Ω or κ_2, approaches zero, these equations can be solved analytically. In practice this can happen if, for example, the magnetic field compensates the contribution of the light shift so that $\Omega_0 = -\Omega_2$. The solution of the system (1) for this special case will be discussed in detail elsewhere. Here we restrict our further discussion to an even simpler model where $\kappa_2 = 0$ and $\Omega = 0$ simultaneously. Although it is not trivial to compensate for the difference in the refractive indices for x- or y-type polarized probe propagating in the spin polarized atomic ensemble, it is very appealing to consider this particular case, due to the following two reasons. 1) There are no quantum correlations under the above conditions induced between the field and atomic subsystems due to regular spin precession and due to average birefringence of the sample. 2) Such an approximation lets us clearly identify the basic features of the macroscopic spin polariton dynamics and their potential importance for the respective quantum information channel. In addition, as a technical simplification, we will assume that the propagation time through the sample is negligibly small, and the retardation effects are unimportant in the frequency domain we are going to discuss.

3. Entanglement Between the Laplace Modes

In an experiment, the light pulse probes an atomic sample of length L during the finite time T. However, because the wave dynamics is developing only in the forward direction in space as well as in time, we can formally extend this process up to infinite interaction time and consider the probe light propagating in a semi-infinite medium. Then the parameter L can be associated with a selected layer in such a medium located at the coordinate $z = L$ and after interaction time T the state of the system is considered at the moment $t = T$. Such an extension lets us define the following Laplace modes for the field and for the atomic subsystem

$$\hat{\Xi}_i^{\rm out}(s) = \hat{\Xi}_i(L,s) = \int_0^\infty dt\, e^{-st}\, \hat{\Xi}_i(L,t)$$

$$\hat{\mathcal{J}}_\mu^{\rm out}(p) = \hat{\mathcal{J}}_\mu(p,T) = \int_0^\infty dz\, e^{-pz}\, \hat{\mathcal{J}}_\mu(z,T) \qquad (2)$$

for $i = 1, 2$ and $\mu = z, y$ respectively. In the above transformations we consider the outgoing field operators at point $z = L$ at any time and define the Laplace s-mode for them. The outgoing atomic operators are considered at a selected moment of time $t = T$ but for an arbitrary spatial location and are described in terms of the spatial Laplace p-mode.

The solution of the system (1) can be conveniently rewritten using relations between the Laplace modes in the following form

$$
\begin{aligned}
e^{-p(s)L}\,\hat{\Xi}_1^{\rm out}(s) \;-\; \frac{2\beta\bar{\Xi}_3}{s}\,e^{-sT}\,\hat{\mathcal{J}}_z^{\rm out}(p(s)) = \\
\int_T^\infty dt\,e^{-st}\,\hat{\Xi}_1^{\rm in}(t) \;-\; \frac{2\beta\bar{\Xi}_3}{s}\int_L^\infty dz\,e^{-p(s)z}\,\hat{\mathcal{J}}_z^{\rm in}(z) \\
e^{-p(s)L}\,\hat{\Xi}_2^{\rm out}(s) \;+\; \frac{2\epsilon\bar{\Xi}_3}{s}\,e^{-sT}\,\hat{\mathcal{J}}_y^{\rm out}(p(s)) = \\
\int_T^\infty dt\,e^{-st}\,\hat{\Xi}_2^{\rm in}(t) \;+\; \frac{2\epsilon\bar{\Xi}_3}{s}\int_L^\infty dz\,e^{-p(s)z}\,\hat{\mathcal{J}}_y^{\rm in}(z) \qquad (3)
\end{aligned}
$$

The initial Heisenberg operators $\hat{\Xi}_i^{\rm in}(t)$ and $\hat{\mathcal{J}}_\mu^{\rm in}(z)$ of the field and atomic subsystems contribute on the right hand side and are defined in their original form as functions on time and spatial coordinates. There are two additional relations, not shown here, which together with the relations (3) lead to the full solution in the Laplace form. The main feature of the solution (3) is that s- and p- Laplace modes are not independent of one another but are coupled by the dispersion relation caused by the wave nature of the process

$$
\begin{aligned}
p = p(s) \;&=\; \frac{A}{s}, \\
A \;&=\; -2\beta\,\epsilon\,\bar{\Xi}_3\,\bar{\mathcal{J}}_x \qquad (4)
\end{aligned}
$$

and because of complete symmetry of the problem, a similar dispersion relation can be written for $s = s(p)$.

As clearly seen from the solution (3) in an extended medium $L \to \infty$ and after long interaction time $T \to \infty$ there will be entanglement developed between the temporal and spatial dynamics inside the polariton wave mode. When the field and atomic subsystems are separated, the mode entanglement would manifest itself in either entanglement or swapping of the outgoing quantum states. The wave nature of such quantum correlations will be visualized after a certain spectral selection in the light subsystem and a spatial spectral selection in the spin subsystem are made.

The important parameter which determines the type of the output quantum correlations is the polariton group velocity. To define this velocity the inverse Laplace transformation should be written in terms of a Fourier integral, with the following parametrization of the Laplace modes: $s = -i\omega$ and $p = iq$. Then the transport dynamics of the correlation wave is characterized by the group velocity

$$
v_g \;=\; \frac{d\omega}{dq} \;=\; -\frac{A}{q^2} \qquad (5)
$$

As one can see the group velocity can be either positive or negative. The latter case appears when the product $\beta\epsilon < 0$, as, in example of alkali atoms, can occur for the probe in the blue wing of the D_2-line. For such a specific situation there is no stationary point in the corresponding wave–packet expansion. The quantum fluctuation are exponentially enhanced in space as well as in time with preserving their quantum correlations. This case requires a special discussion. Below, we consider an alternative, more typical situation for the wave dynamics, when the group velocity is positive and $\beta\epsilon > 0$. In the case of alkali atoms this relates to the red wing of D_2-line.

4. Quantum Memory and Readout Protocols

Consider an experimental situation when a low-frequency mode with the frequency ω, such that $\omega T < 1$, is detected in the output polarization state of the transmitted light. For this mode, according to the dispersion law (4) and (5), the group velocity can be quite low. As a consequence, the input quantum state of the entire system will be transported through the sample as a slow propagating wavepacket. Then the output Heisenberg operators of the light subsystem will be actually formed in the readout process of the input quantum operators of atomic spins and preferably of those spatial components which are distributed near respective high-frequency spatial mode $q = A/\omega$.

This can be demonstrated by the graphs plotted in figure 1. All the dependencies are the expectation variances for the following target integral observables

$$\hat{\Xi}_i^{\rm out} = \int_0^T dt\, \cos(\omega t)\, \hat{\Xi}_i^{\rm out}(t) \tag{6}$$

for $i = 1, 2$. The calculations are performed for the red wing of D_2-line of ^{133}Cs, for detuning of the probe about -1200 MHz from the $F_0 = 4 \to F = 5$ hyperfine transition. The product of parameters $-ALT$ is varied from 0 (left border of the graph) to the level of 2 (right border of the graph) so that our chosen frequency of detection, $\omega T = 0.5$, is asymptotically coupled with the spatial mode with the wave number $qL = 4$, in the way explained in the previous section. The input quantum states of the light and spin subsystems obey Poissonian statistics. The blue curve shows how the input quantum fluctuations of the Stokes components are mapped onto the output state. As one can see for the extended medium with high optical activity βJ (where $J = \bar{\mathcal{J}}_x L$ is the total angular momentum of the sample) the contribution of the input field fluctuations becomes negligible. The main impact on the output variances is made by the input fluctuations of the spin subsystem. It also follows from the graphs that the variance of the $\hat{\Xi}_1^{\rm out}$ component is much bigger than $\hat{\Xi}_2^{\rm out}$ and the latter is below the standard quantum deviation. This is a direct consequence of inequality $\beta \gg \epsilon$, which is typical in the case of the D_2 transition of alkali atoms. So the mapping of the state of light is in this case combined with the squeezing operation on the state. The output polarization state of the probe light in the low-frequency domain of its fluctuation spectrum becomes squeezed and the high-frequency spatial modes of the spin states $\hat{J}_z^{\rm in}$ and $\hat{J}_y^{\rm in}$ are respectively mapped into the integral modes of the $\hat{\Xi}_1^{\rm out}$ and $\hat{\Xi}_2^{\rm out}$ Stokes components.

Let us reverse the experimental situation and let the quantum state of the spin subsystem be controlled by the high-frequency polarization fluctuations of light, such that $\omega T > 1$. Then the quantum memory protocol can be achieved and the quantum state of

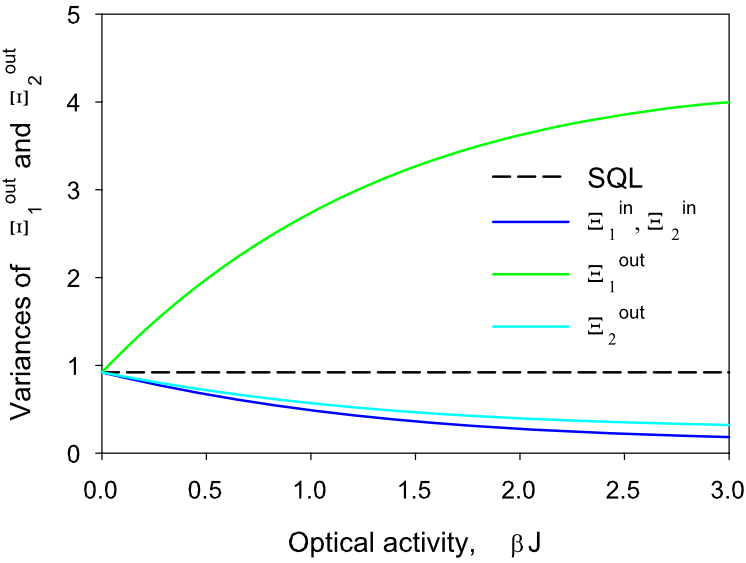


Figure 1. The variances of the integral output Stokes components, defined by Eq.(6) for $\omega T = 0.5$, as a function of the optical activity of the sample βJ. The dotted curve indicates the vacuum noise (standard) quantum limit (SQL). The blue curve indicates the contribution of the input Stokes components.

these modes can be mapped into the output low-frequency spatial modes of the spin subsystem. Indeed, in this case the transfer of input correlations is fast because the polariton group velocity v_g is high enough. The original input quantum state of the system will be transported out of the sample with the polariton wave. Then the low-frequency spatial modes (with $qL < 1$) of the output spin state are mainly formed via accumulation of the input quantum fluctuations of the transmitted light. As a result of the complete symmetry of the equations (1) with respect to the light and spin subsystems, the memory protocol can be visualized by the same graphs shown in figure 1 with the following change in notation. The target observables should be now associated with the spin subsystem

$$\hat{J}_{\mu}^{\mathrm{out}} = \int_0^L dz\, \cos(qz)\, \hat{\mathcal{J}}_{\mu}^{\mathrm{out}}(z) \tag{7}$$

for $\mu = z, y$. The blue curve indicates the reduction of the input state mapping onto these observables. The abscissa of the plot should be associated with $\beta \bar{\Xi}_3 T$, which is the angle of the collective spin rotation if the probe light were circular polarized. After the change in notation the graphs of figure 1 show how the initial polarization quantum state of light, which existed in the high-frequency domain of its fluctuation spectrum, can be mapped into the integral output quantum state of atomic spins. Asymptotically (for the parameters used in our calculations) this results in the mode coupling between $\omega T = 4$ time mode of the Stokes variables and the $qL = 0.5$ spatial mode of atomic spins.

5. Conclusion

In this report we have discussed how interchange between the polarization quantum states of light and atomic spins could be implemented via the polarization sensitive interaction between these subsystems. The main result is the identification of the spectral domains, where the quantum states of the input temporal or spatial fluctuations could be transferred from one subsystem to another. If the polarization fluctuations of the light and spin subsystems were to carry quantum information it would be possible to create quantum information protocols for writing in or reading out this information. This specific quantum channel does not exactly copy an original quantum state from one physical carrier to another, but creates a new quantum state in the target system related in a known way to the quantum state originally existing in the source system. This can be done because of the multimode nature of interaction in the entire atoms-field system and the protocols are related only to the specially selected spectral domains. In a general case the input-output transformations in the usual space-time representation are expressed via the fundamental solution of the wave-type Heisenberg equations (1) in integral form, see [4]. For interaction limited in space as well as in time, the idealized mode description can only approximate the real interaction process. Because of the multi-mode character of the input-output transformation the fidelity, usually applied to describe the quality of a quantum information protocol for relatively simple systems, has to be revised in order to become applicable. This work is in progress and will be published elsewhere.

Acknowledgements

The work was supported by the Russian Foundation for Basic Research (RFBR-05-02-16172-a) and by the European grant within network COVAQIAL. O.S.M. would like to acknowledge the financial support from the charity Foundation "Dinastia".

References

[1] Samuel L. Braunstein abd Peter van Loock, Quantum information with continuous variables, *Rev. Mod. Phys.* **77** (2005), 513.
[2] B. Julsgaard, A. Kozhekin, E.S. Polzik, Experimental long-lived entanglement of two macroscopic objects, *Nature* **413** (2001), 400.
[3] B. Julsgaard, J. Sherson, J. Fiurasek, J.I. Cirac, and E.S. Polzik, Experimental demonstration of quantum memory for light, *Nature* **432** (2004), 482.
[4] D.V. Kupriyanov, O.S. Mishina, I.M. Sokolov, B. Julsgaard, E.S. Polzik, Multimode entanglement of light and atomic ensembles via off-resonant coherent forward scattering, *Phys. Rev. A* **77** (2005), 032348.

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

From Network Complexity to Time Complexity Via Optimal Control[1]

Thomas Schulte-Herbrüggen [a,2], Andreas Spörl [a], Navin Khaneja [b] and Steffen Glaser [a]

[a] *Department of Chemistry, Technical University Munich, 85747 Garching, Germany*
[b] *Division of Applied Sciences, Harvard University, Boston, USA*

Abstract. We demonstrate the substantial speed-up of quantum compilers directly translating quantum modules into the machine code of evolutions under available controls and drift of the pertinent experimental hardware. This is exemplified in two settings of time scales, where (i) the non-local coupling evolution is limiting (as in NMR) and (ii) the local controls take about twice the time of the non-local ones (as in coupled Josephson charge qubits). Finally we give a mini-tutorial proving Pontryagin's Maximum Principle in a quantum setting.

Keywords. Timeoptimal quantum control, quantum compiler, complexity.

1. Controllable Hardware for Quantum Computation and Quantum Simulation

In view of quantum computation and quantum simulation, recent years have seen an increasing array of quantum systems that can be coherently controlled. These include pure-state systems as well as ensembles [1]. For an n-qubit system to be fully controllable [2], all the qubits have to be selectively addressable and coherently coupled in an arbitrary connected coupling topology [3]. Then the entire Lie algebra $\mathfrak{su}(2^n)$ can be generated by the available controls and the drift term, in other words, local and universal quantum gates can be implemented. Such a system can be regarded as a hardware candidate for a quantum computer, if it also fulfills the other DiVincenzo criteria [4].

2. Compiling by Optimal Control: from Quantum Gates to Machine Code

A quantum compiler [5,6,7,8] is usually seen as translating the language of quantum modules into a sequence of universal two-qubit gates (CNOT, controlled phase etc.) and local operations [9]. However, here we advocate to use optimal control for directly going into the 'machine code' of controls in the experimental harware setting. This means decomposing a quantum module U_G into a timeoptimal (if not relaxation-optimised) sequence of evolutions $U_G = U_M U_{(M-1)} \cdots U_1$ under the system Hamiltonian (drift) and experimentally available controls.

[1] Here we also try to continue a fruitful discussion on *quantum compilers* with the audience at the 2005-ASI.

[2] Correspondence to: Thomas Schulte-Herbrüggen, Dept. Chemistry, TU-Munich, Lichtenbergstrasse 4, 85747 Garching, Germany. Tel.: +49 89 289 13312; Fax: +49 89 289 13210; E-mail: tosh@ch.tum.de

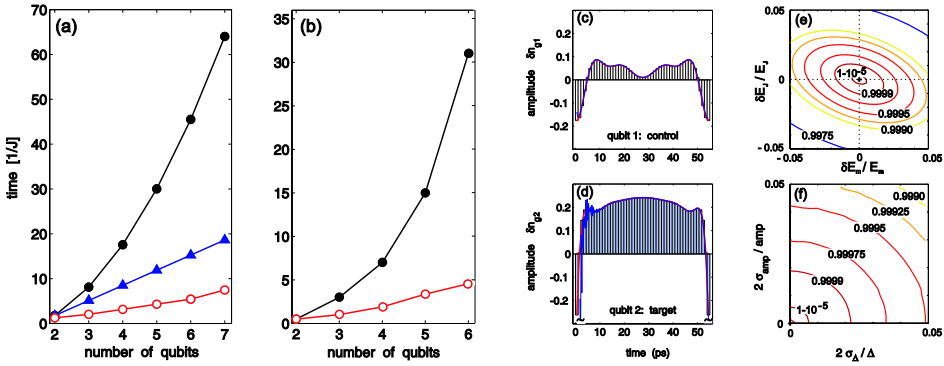


Figure 1. (a) Time complexity of the QFT in linear coupling topologies. Upper traces give analytical times associated with standard-gate decompositions (•) [15] and optimised scalable gate decompositions (▲) [16]. Lowest trace: speed-up by time-optimal control with shortest numerical realisations obtained (○) rounded to $0.01\,J^{-1}$. (b) Time complexity of the C^{n-1}NOT-gate on complete coupling topologies. Upper trace: analytical times for decomposition into standard gates (•) [17]. Lower trace: speed-up by time-optimal control with shortest times (○) currently needed for realising C^{n-1}NOT by numerical control rounded to $0.01\,J^{-1}$. (c,d) Fastest charge level controls obtained for realising a CNOT-gate on two coupled charge qubits (c: control qubit, d: working qubit). The total gate charges for the qubits are $n_{g\nu} = n_{g\nu}^{(0)} + \delta n_{g\nu}$ with $\nu = 1, 2$. $n_{g1}^{(0)} = 0.24$, $n_{g2}^{(0)} = 0.26$ and the qubit energies were taken from the experimental values in [18]. The 50 piecewise constant controls are shown as bars (uniform width $\Delta = 1.1$ ps); the trace fidelity is $\frac{1}{2^N}\left|\mathrm{tr}\{U_{\mathrm{target}}^{\dagger} U_T\}\right| > 1-10^{-9}$. Red lines give harmonic fits, while the blue ones superimposed show a pulse synthesised by an LCR-filter. (e) Trace fidelities resulting from the controls of (c,d) when the parameters E_m (coupling) and E_J (tunneling energy) vary by $\pm 5\%$. (f) Fidelities under Gaussian noise on control amplitudes and time intervals Δ parameterised by the standard deviations $2\sigma_\Delta/\Delta$ and $2\sigma_{\mathrm{amp}}/\mathrm{amp}$ ranging from 0 to 5%.

3. Form Network Complexity Towards Time Complexity

The number of elementary gates required for implementing a quantum module defines the network complexity, which however, often translates into too coarse an estimate for the time required to implement a quantum module [10,11,12] in a concrete experimental setting: there is no simple one-to-one relation between the two complexities, because of different time-scales between local and non-local controls, different coupling topologies allowing for different degrees of parallelisation or requiring indirect realisation, and different types of coupling interactions matching different sets of elementary gates.

Thus we leave the usual approach of decomposing gates into sets of discrete universal building blocks by exploiting the differential geometry of the unitary group for quantum control [3,13] in order to obtain constructive bounds to minimal time both as close to the experimental setting and as tight as possible [14]. In the limit of zero cost for the fast local controls (as in NMR) compared to the slow coupling interactions, we give decompositions for the Quantum Fourier Transform (QFT) and the multiply-controlled NOT-gate C^{n-1}NOT that are dramatically faster than the fastest compilations into standard gates known so far as seen in Figs. 1 (a) and (b). In contrast, for coupled Josephson charge qubits the local controls are time-limiting. Here the controls of Fig. 1 (c) and (d) are 5 times faster than in Ref. [18], while being fairly robust against 5% variation of tunneling and coupling energy (Fig. 1 (e)) as well as against Gaussian noise on the control amplitudes and unit times (Fig. 1 (f)). Details can be found in Refs. [14,19].

4. Principles of Quantum Control—a Short Tutorial

The goal is to prove the corner stone of control theory, Pontryagin's celebrated Maximum Principle [20] in a quantum setting [21]. It allows for optimising a scalar quality functional $L(t)$ subject to the operator eqn. of motion $\dot{X} = F(X, u, t)$ with u as the vector of control amplitudes $u = \big(u(t_1), \ldots, u(t_k), \ldots, u(t_M)\big)$. The concept can readily be used to steer quantum states (density operators) following an eqn. of motion $\dot{X} = -i[H, X]$, or to generate unitary propagators: subject to $\dot{U} = -iHU$ one may maximise the trace-fidelity by choosing the quality functional to be $L(T) = |\mathrm{tr}\{U_{\mathrm{gate}}^{\dagger} U(T)\}|$ [14].

The underlying variational calculus makes use of the unitary group viewed as a smooth manifold. Let M be a complete normed space, *i.e.* Banach space, with a manifold structure allowing for k-times differentiable curves. So it denotes a C^k-Banach manifold. Then the tangent space TM_x to M at point $x \in M$ is the set of all tangent vectors at x. TM_x is a topological vector field, and—if M carries the structure of a Lie group—$TM_{\mathbb{1}}$ is the Lie algebra of M. The differential structure of Banach manifolds is locally expressed in terms of Fréchet derivatives to be introduced next. The notion of the Fréchet-differential coincides with the tangent map in manifolds or Lie algebras $F'(x) : T\,M_x \to T\,N_{F(x)}$ in local coordinates [22].

Definition 1 (Fréchet Derivative) *Let the operator $F : M \to N$ represent a mapping between the Banach manifolds M and N and let F be defined on some open neighbourhood to $X_0 \in M$ of radius r denoted $U(X_0) = \{X \in M |\ \|X - X_0\| < r\}$. Then F has a Fréchet differential $F'(X_0)$ in X_0 if for all $H \in M$ with $\|H\| < r$ there is an operator $\varepsilon_H \to 0_N \in N$ if $H \to 0_M \in M$ and a continuous linearisation $F'(X_0)$ fulfilling*

$$F(X_0 + H) - F(X_0) = F'(X_0)H + \varepsilon_H \|H\| \ .$$

In the complex domain, the Fréchet differential $dF(X_0,\, H) \ = \ F'(X_0)H$ avoids the Cauchy-Riemann conditions, while in the real case it relates to the Jacobian. Operator differential calculus on Banach algebras with a Fréchet differentiable manifold structure follows—of course—the chain and the product rule as well as it allows to compute the differential of inverse mappings.

Theorem 1 (Pontryagin's Maximum Principle) *Consider a (quantum) system governed by the (operator) eqn. of motion $\dot{X}(t) = F(X, u, t)$ with control amplitudes u. For $u_*(t)$ to be an optimal control steering $X(0)$ into some $X(T)$ leading to critical points of the scalar quality functional $J[X(t)] = \int_0^T dt\ L(X(t))$ over the entire time course, it is sufficient that there is*

1. *an adjoint system $\lambda(t)$ satisfying the equation of motion $\dot{\lambda} = -\frac{\partial h}{\partial X}$ by virtue of*
2. *a scalar Hamiltonian function $h(X, \lambda, u, t) := L(X, u, t) + \langle \lambda(t) | F(X, u, t) \rangle$, so $\dot{X}(t) \equiv F(X, u, t) = \frac{\partial h}{\partial \lambda^{\dagger}}$, where (i) h attains critical points for optimal controls $u_*(t)$, i.e., $\frac{\partial h}{\partial u_*(t)} = 0$ at almost all $0 \le t \le T$ and (ii) $h(T) = 0$.*

Proof: The Fréchet derivative provides a matrix $\frac{\partial L}{\partial X}$ and a vector $\frac{\partial L}{\partial u}$. Thus for $J[X(t)] = \int_0^T dt\, L(X, u, t)$ one gets the first variation in X and u

$$\delta J \stackrel{1^{\circ}}{=} J(X + \delta X, u + \delta u, t) - J(X, u, t) = \int_0^T dt \{ \langle \frac{\partial L}{\partial X} \, |\delta X\rangle + \langle \frac{\partial L}{\partial u} \, |\delta u\rangle \} + L(t)\delta t \, \Big|_0^T \quad ,$$

where δX depends on the variation of the control δu following $\dot{X} = F(X,u,t)$ accounted for by an operator-valued Lagrange multiplier $\lambda(t)$ associated with the zero-cost term $J_\lambda := \int_0^T dt\langle\lambda(t)|F(X,u,t) - \dot{X}\rangle = 0$. Its variation in X and u reads

$$\delta J_\lambda \stackrel{1^\circ}{=} \int_0^T dt\Big\{\langle\frac{\partial\langle\lambda|F\rangle}{\partial X}\ |\delta X\rangle + \langle\frac{\partial\langle\lambda|F\rangle}{\partial u}\ |\delta u\rangle + \langle\dot{\lambda}|\delta X\rangle\Big\} - \langle\lambda|\delta X\rangle\ \Big|_0^T \quad .$$

Sorting terms one gets for the total of the first variations (with $L(0) = 0$ and $\delta X(0) = 0$)

$$\delta J + \delta J_\lambda = \int_0^T dt\Big\{\langle\frac{\partial L + \partial\langle\lambda|F\rangle}{\partial X}\ + \dot{\lambda}|\delta X\rangle + \langle\frac{\partial L + \partial\langle\lambda|F\rangle}{\partial u}\ |\delta u\rangle\Big\} + L(T)\delta t + \langle\lambda(T)|F(T)\rangle\delta t.$$

The final condition $X(T+\delta t) + \delta X(T+\delta t) = X(T)$ implies to first order $\dot{X}(T)\delta t + \delta X(T) = 0$, so $\delta X(T) = -\dot{X}(T)\delta t = -F(T)\delta t$; i.e. $-\langle\lambda(T)|\delta X(T)\rangle = \langle\lambda(T)|F(T)\rangle\delta t$. Introducing the scalar-valued hamiltonian function

$$h(X,\lambda,u,t) := L(X,u,t) + \langle\lambda(t)|F(X,u,t)\rangle = L + \langle\lambda|\dot{X}\rangle$$

one finally arrives at

$$\delta J + \delta J_\lambda = \int_0^T dt\Big\{\langle\frac{\partial h}{\partial X}\ + \dot{\lambda}|\delta X\rangle + \langle\frac{\partial h}{\partial u}\ |\delta u\rangle\Big\} + h(X,\lambda,u,t=T)\delta t\ .$$

Controls $u_*(t)$ leading to quality-optimising trajectories $X_*(t)$ and their adjoints $\lambda_*(t)$ result if

$$\begin{aligned}
\dot{\lambda}_*(t) &= -\frac{\partial h(X_*,\lambda_*,u_*,t)}{\partial X_*} \\
\dot{X}_*(t) \equiv F(X_*,u_*,t) &= \frac{\partial h(X_*,\lambda_*,u_*,t)}{\partial \lambda_*} \\
\frac{\partial h(X_*,\lambda_*,u_*,t)}{\partial u_*} &= 0 \\
h(X_*,\lambda_*,u_*,T) &= 0 \quad ,
\end{aligned}$$

as stated in Pontryagin's Maximum Principle. ■

Remark 1 *Note that* $\dot{X} = \frac{\partial h}{\partial \lambda}$ *and* $\dot{\lambda} = -\frac{\partial h}{\partial X}$ *constitute a Hamiltonian system, as do—by virtue of Ehrenfest's theorem—the expectation values* $\frac{d}{dt}\langle Q_j\rangle = i\langle[H,Q_j]\rangle$ *and* $\frac{d}{dt}\langle P_j\rangle = i\langle[H,P_j]\rangle$ *e.g. in the textbook case of a spin-less particle of mass* m *in a stationary potential* $V(Q)$ *with* $H(P,Q) = \frac{1}{2m}P^2 + V(Q)$*, where* $\frac{d}{dt}\langle Q_j\rangle = \langle\frac{\partial H}{\partial P_j}\ \rangle = \frac{1}{m}\langle P_j\rangle$ *and* $\frac{d}{dt}\langle P_j\rangle = -\langle\frac{\partial H}{\partial Q_j}\ \rangle = -\langle\nabla V(Q_j)\rangle$ *see e.g. [23].*

Acknowledgements

Fruitful discussion with the participants at the ASI—in a most perfect setting provided by the organisers—is gratefully acknowledged. Ref. [19] is a collaboration with Frank Wilhelm's group at LMU Munich and Ville Bergholm at HUT Espoo. This work was supported by DFG through the grant Gl 203/4-2 in the incentive SPP 1078.

References

[1] for references: T. Schulte-Herbrüggen, A. Spörl, R. Marx, N. Khaneja, J. Myers, A. Fahmy, and S.J. Glaser: Quantum Computing Implemented via Optimal Control, in: *Lecture Notes on Quantum Information*, D. Bruss and G. Leuchs, (Eds.), Wiley, Berlin, in press.

[2] F. Albertini and D. D'Alessandro: Notions of Controllability for Bilinear Multilevel Quantum Systems, *IEEE Trans. Automat. Control* **48** (2003), 1399–1403.

[3] S. J. Glaser, T. Schulte-Herbrüggen, M. Sieveking, O. Schedletzky, N. C. Nielsen, O. W. Sørensen, and C. Griesinger: Unitary Control in Quantum Ensembles: Maximising Signal Intensity in Coherent Spectroscopy, *Science* **280** (1998), 421-424.

[4] D. P. DiVincenzo: The Physical Implementation of Quantum Computation, *Fortschr. Phys.* **48** (2000), 771–783.

[5] R. R. Tucci: A Rudimentary Quantum Compiler, quant-ph/9902062 (1999).

[6] N. Khaneja, S. J. Glaser, and R. Brockett: Sub-Riemannian Geometry and Time-Optimal Control of Three-Spin Systems: Quantum Gates and Coherence Transfer, *Phys. Rev. A* **65** (2002) 032301.

[7] M. A. Nielsen, quant-ph/0502070 (2005).

[8] H. N. Sá-Earp and J. K. Pachos, quant-ph/0505128 (2005).

[9] D. Deutsch: Quantum Theory, the Church-Turing Principle, and the Universal Quantum Computer, *Proc. Royal Soc. London A* **400** (1985), 97–117.

[10] G. Vidal, K. Hammerer, and J. I. Cirac: Interaction Cost of Nonlocal Gates, *Phys. Rev. Lett.* **88** (2002), 237902.

[11] P. Wocjan, D. Janzing, and T. Beth: Simulating Arbitrary Pair-Interactions by a Given Hamiltonian: Graph-Theoretical Bounds on the Time-Complexity, *Quant. Inf. Comput.* **2** (2002), 117–132.

[12] A. M. Childs, H. L. Haselgrove and M. A. Nielsen: Lower Bounds on the Complexity of Simulating Quantum Gates, *Phys. Rev. A* **68** (2003), 052311.

[13] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser: Optimal Control of Coupled Spin Dynamics: Design of NMR Pulse Sequences by Gradient Ascent Algorithms, *J. Magn. Reson.*, **172** (2005), 296–305.

[14] T. Schulte-Herbrüggen, A. K. Spörl, N. Khaneja, and S. J. Glaser: Optimal Control-Based Efficient Synthesis of Building Blocks of Quantum Algorithms: A Perspective from Network Complexity towards Time Complexity, *Phys. Rev. A* **72** (2005), 042331.

[15] A. Saito, K. Kioi, Y. Akagi, N. Hashizume, and K. Ohta: Actual Computational Time-Cost of the Quantum Fourier Transform in a Quantum Computer using Nuclear Spins, quant-ph/0001113 (200).

[16] A. Blais: Quantum Network Optimisation, *Phys. Rev. A* **64** (2001), 022312.

[17] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. W. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter: Elementary Gates for Quantum Computation, *Phys. Rev. A* **52** (1995), 3457–3467.

[18] T. Yamamoto, Yu. A. Pashkin, O. Astafiev, Y. Nakamura, and J. S. Tsai: Demonstration of Conditional Gate Operation Using Superconducting Charge Qubits, *Nature* (London) **425** (2003), 941–944.

[19] A. K. Spörl, T. Schulte-Herbrüggen, S. J. Glaser, V. Bergholm, M. Storcz, J. Ferber, and F. K. Wilhelm: Optimal Control of Coupled Josephson Qubits, quant-ph/0504202 (2005).

[20] L. S. Pontryagin, V. G. Bol'tanskii, R. S. Gamkrelidze, and E. F. Mischenko: *The Mathematical Theory of Optimal Processes*, Pergamon Press, New York, 1964.

[21] A. G. Butkovskiy and Yu. I. Samoilenko: *Control of Quantum-Mechanical Processes and Systems*, Kluwer, Dordrecht, 1990.

[22] E. Zeidler: *Non-Linear Functional Analysis and its Applications* Vol. 1 chapter 4 and Vol. 4 chapters 73 and 74. Springer, New York, 1986 (Vol. 1) and 1988 (Vol. 4).

[23] L. D. Landau and E. M. Lifshitz: *Theoretical Physics, Quantum Mechanics*, Pergamon, 1977.

This page intentionally left blank

Quantum Information Processing
D.G. Angelakis et al. (Eds.)
IOS Press, 2006
© 2006 IOS Press. All rights reserved.

Author Index

Augusiak, R.	19	Kobayashi, N.	312
Bagayev, S.N.	341	Kulik, S.	281
Barnes, C.H.W.	321	Kupriyanov, D.V.	346
Bombin, H.	30	Kwek, L.-C.	83
Bose, S.	218	Lebedev, V.N.	341
Bruß, D.	3	Martin-Delgado, M.A.	30, 191
Bukach, A.A.	326	Mateus, P.	180
Burgarth, D.	218	Mekhov, I.B.	341
Chatzisavvas, K.Ch.	330	Mendonça, J.T.	337
Chekhonin, I.A.	341	Mishina, O.S.	346
Chen, J.	83	Mizubayashi, J.	312
Cheung, D.	93	Moroshkin, P.V.	341
Cirac, I.	191	Mosca, M.	121
Daskaloyannis, C.	330	Murakami, K.	312
Doxaras, I.N.	170	Oh, C.H.	83
Egorov, V.S.	341	Omar, Y.	180
Ekert, A.K.	v, 93	Panos, C.P.	330
Ericsson, M.	238	Polzik, E.S.	346
Ferreira, A.	89, 337	Popescu, S.	99
Giovannetti, V.	218	Popp, M.	191
Gisin, N.	99	Raimond, J.-M.	294
Glaser, S.	353	Sanders, B.C.	184
Gottesman, D.	159	Schulte-Herbrüggen, T.	353
Grishanin, B.	113	Scott, A.J.	184
Guerreiro, A.	89, 337	Short, A.J.	99
Hamma, A.	175	Smyrnakis, I.	246
Haruyama, J.	312	Spörl, A.	353
Herbauts, I.	242	Sych, D.	113
Holevo, A.S.	41	Terno, D.	61
Horodecki, P.	19	Travaglione, B.C.	93
Hübel, H.	255	Vedral, V.	89, 337
Ioannou, L.M.	93	Verstraete, F.	191
Ionicioiu, R.	175	Vlasov, A.Yu.	103
Jozsa, R.	137	Walgate, J.	109, 184
Kaszlikowski, D.	83	Wu, C.	83
Kay, A.	238	Zadkov, V.	113
Khaneja, N.	353	Zanardi, P.	175
Kilin, S.Ya.	326	Zeilinger, A.	255

This page intentionally left blank

This page intentionally left blank

This page intentionally left blank